Channel Hardening of IRS-Aided Multi-Antenna Systems: How Should IRSs Scale?

Ali Bereyhi, *Member, IEEE*, Saba Asaad, *Member, IEEE*, Chongjun Ouyang, *Student Member, IEEE*, Ralf R. Müller, *Senior Member, IEEE*, Rafael F. Schaefer, *Senior Member, IEEE* and H. Vincent Poor, *Life Fellow, IEEE*

Abstract— Unlike active array antennas, intelligent reflecting surfaces (IRSs) are efficiently implemented at large dimensions. This allows for traceable realizations of large-scale *IRS-aided* MIMO systems in which not necessarily the array antennas, but the passive IRSs are large. It is widely believed that large IRS-aided MIMO settings maintain the fundamental features of massive MIMO systems, and hence they are the implementationally feasible technology for establishing the performance of large-scale MIMO settings. This work gives a rigorous proof to this belief. We show that using a large passive IRS, the end-to-end MIMO channel between the transmitter and the receiver always hardens, even if the IRS elements are strongly correlated.

For the fading direct and reflection links between the transmitter and the receiver, our derivations demonstrate that as the number of IRS elements grows large, the capacity of end-to-end channel converges in distribution to a real-valued Gaussian random variable whose variance goes to zero. The order of this drop depends on how the physical dimensions of the IRS grow. We derive this order explicitly. Numerical experiments depict that the analytical asymptotic distribution almost perfectly matches the histogram of the capacity, even in practical scenarios.

As a sample application of the results, we use the asymptotic characterization to study the dimensional trade-off between the transmitter and the IRS. The result is intuitive: For a given target performance, the larger the IRS is, the less transmit antennas are required to achieve the target. For an arbitrary ergodic and outage performance, we characterize this trade-off analytically. Our investigations demonstrate that using a practical IRS size, the target performance can be achieved with significantly small end-to-end MIMO dimensions.

Index Terms—Reconfigurable intelligent surface, asymptotic channel hardening, large-system analysis, ergodic capacity, outage probability.

I. Introduction

Channel hardening is a property of massive multiple-input multiple-output (MIMO) systems indicating that the random fading process in a MIMO channel becomes a deterministic effective path-loss, as the dimensions of the MIMO channel grow large at least at one side; see for instance [1]–[3]. From the information-theoretic point of view, this is the key property of massive MIMO systems which leads to their significant performance gains in various aspects. These gains are rather well-studied in the literature; e.g., [4]–[6].

Despite the promising performance, massive MIMO systems are still considered to be implementationally intractable. This follows from the fact that deploying large active antenna arrays imposes high hardware cost and energy consumption, as well as unrealistic design requirements [7]–[9]. The issue becomes more crucial in practice, since several recent studies show that the fundamental features of massive MIMO systems[1] do not hold with practical dimensions in several settings, e.g., MIMO systems with highly-correlated channels, scenarios with small scattering angular spread and cell-free networks; some detailed studies can be found in [10] and [11].

Intelligent reflecting surface (IRS)-aided MIMO communication has been recently proposed as an implementationally-tractable solution to realize large-dimensional MIMO systems [12]–[14]. An IRS is a two-dimensional surface, with a large number of reconfigurable passive reflective elements. These elements are capable of tuning the wireless propagation dynamically to achieve various design objectives: For instance, in indoor millimeter-wave applications, by installing IRSs out of the transmission site, e.g., on interior walls or room ceilings, the blockage issue of the millimeter-wave communication can be reduced significantly, due to the link established between the transmitter and the users through the IRSs [15]. As other examples of design targets which are achieved by employing IRSs in the system, we can name spatial interference suppression, accurate three-dimensional beamforming, and providing a more favorable propagation environment [16]–[19].

As mentioned, unlike classical large-scale MIMO technologies which require large active arrays of antennas, large IRS-aided MIMO systems, i.e., MIMO settings with large reflecting surfaces, are implementationally feasible. This follows the fact that IRSs *passively* reflect the incoming signals and do not employ radio frequency chains. As a result, the hardware cost and energy consumption at large scale decrease notably [20], [21]. Inspired by this appealing advantages, IRSs are employed to a develop practically-tractable alternative designs for several large-scale MIMO technologies, such as massive MIMO [22], [23], cognitive radio [24], [25], non-orthogonal multiple access (NOMA) [26], [27] and simultaneous wireless information and power transform systems [28], [29]. The key motivation behind these designs comes from this widely-accepted belief: *The key features of a large-scale MIMO technology can be achieved via a small-dimension MIMO setting assisted by large IRSs.* Although under idealistic assumptions this belief seems to be intuitive, it is not easy to conclude its validity in practical IRS-

---

[1]These features are mainly described by the channel hardening.

aided settings with highly-correlated IRS elements and multiple propagation links between the transmitter and the receiver. This work aims to give a rigorous proof for this believe by considering the fundamental channel hardening property in a practical IRS-aided multi-antenna setting.

A. Main Objective and Contributions

This work gives an answer to this intriguing question: *How does the end-to-end channel in IRS-aided multi-antenna systems harden, when only the IRS size grows large?* We answer this question by deriving the asymptotic distribution of the channel capacity expression in an IRS-aided fading multiple-input single-output (MISO) system[1] whose number of transmit antennas is fixed and rather small, when the number of IRS elements grow asymptotically *large*. Our derivations demonstrate that for any IRS covariance matrix, the channel capacity converges in distribution to a real-valued Gaussian distribution whose mean increases unboundedly large, and whose variance tends to zero, as the IRS size grows asymptotically large.

The analytical derivations of this study gives the following answer to our target question: *Even with a strongly-correlated IRS, the end-to-end channel hardens regardless of the transmit array size[2]*. Interestingly, this property is achieved *without* any need for frequent tuning of IRS elements.

In addition to the main results, this study has several other contributions which are briefly highlighted below:

- We characterize the distribution of the channel capacity for a generic IRS-aided fading MISO setting in which both direct and reflecting links are available between the transmitter and receiver. We further take into account the impact of line-of-sight channels and consider an arbitrary scaling of the IRS area and a general correlation among the reflecting elements. To the best of our knowledge, the characterization of channel hardening for such setting has been left open in the literature; see for instance the recent studies in [30] and [31].
- For a given IRS and transmit array architecture, we determine the *exact* distribution of the end-to-end signal-to-noise ratio (SNR) for an arbitrary choice of phase-shifts at the IRS. We give a universal upper-bound for the mean and variance of this distribution, and propose an effective choice of IRS phase-shifts that guarantees the end-to-end channel hardening.
- We validate our analytical derivations through numerical experiments. Our investigations show that the asymptotic distribution fits almost perfectly the histogram of channel capacity for practical IRS sizes[3]. We further conduct several numerical experiments to investigate channel hardening for classical and extreme IRS correlation scenarios.
- To characterize the speed of channel hardening, we derive a lower-bound on the mean and a uniform upper-bound on the variance of the channel capacity. Invoking the bounds,

we show that for any covariance matrix at the IRS whose largest eigenvalue grows *sub-linearly*[4] with the IRS size, the limit of the variance, when the IRS size grows large, is zero. Using extreme-case investigations, we demonstrate that this conclusion is further extended to most extreme scenarios with the largest eigenvalue growing *linearly* and the physical dimension of IRS being fixed.

- As a sample application, we use the main results to investigate the dimensional trade-off between the transmitter and the IRS. The result is intuitive: *By increasing the IRS size, the transmit array size required to achieve a given target performance reduces*. We derive the trade-off curve analytically two performance metrics, namely the ergodic capacity and outage probability.

B. Related Work

Channel hardening in IRS-aided MIMO systems is studied in the literature for some restricted settings: In [32], channel hardening is discussed in an isotropic scattering environment considering an IRS-aided setting with a single-antenna transmitter and receiver, i.e., a single-input single-output (SISO) setting. The analysis is extended to scenarios with multiple IRSs and *uncorrelated Rician* fading in [33]. The concept of channel hardening is further addressed in an alternative way in [34], where the authors show that in an IRS-aided SISO system with generic Nakagami-m fading channels, the end-to-end channel between the transmitter and the receiver, in the presence of a direct link, becomes nearly deterministic. The results of this study further demonstrate that increasing the number of IRS elements, as well as reducing the fading severity, enhances the channel hardening property.

In addition to classical point-to-point settings, channel hardening has been further discussed in alternative scenarios. For instance, channel hardening is investigated in [35] for an IRS-aided SISO NOMA system, where the IRS is assumed to be realized by the intelligent omni-surfaces technology, and the reflecting elements are considered to be correlated in general. The results show that for such settings, the average achievable rate converges asymptotically to that of uncorrelated channels, when the number of IRS elements goes to infinity. Another example is a MIMO setting with a fully stand-alone IRS-based transmitter. For such settings, the authors of [30] show that by growing the size of the transmitter unit, the user channels become deterministic and mutually orthogonal.

Despite connections to the above literature, the most related lines of work to our study are those given in [23], [36] and [31]: Starting with the study in [36], the authors show that the conventional form of channel hardening is not valid for an IRS-aided MIMO setting. The result depicts that by increasing the number of transmit antennas, while keeping the number of IRS elements fixed, the hardening property does not hold. This finding follows from the fact that the channel between the IRS and transmitter is identical for all receivers. As a result, by growth of transmit array size, the randomness of the fading process grows large, and hence the effective end-to-end

---

[1]It is worth mentioning that the results readily extend to cases with multiple receive antennas. A MISO setting is assumed mainly to keep the derivations tractable. This assumption however impacts neither the analytical approach nor the final conclusions of this study.

[2]In fact, it hardens asymptotically with only a single transmit antenna.

[3]And even much smaller sizes of IRS.

[4]The definition of sub-linear growth becomes clear in the forthcoming sections.

channel does not converge to a deterministic channel. In their following work, i.e., [23], the authors invoke the Lindeberg-Feller central theorem [37] to demonstrate that the channel hardening in IRS-aided systems occurs when the *number of IRS elements* grows large[1].

The study in [31] derives the distribution of the input-output mutual information for an IRS-aided MIMO setting in which the communication is carried out *only* through the IRS, i.e., there is no direct link between the transmitter and receiver. The analysis invokes random matrix theory to characterize the setting in an asymptotic regime, in which both the numbers of IRS elements and transmit antennas grow large with a fixed-ratio. The authors further utilize the results to derive a closed-form expression for the outage probability and the optimal values of IRS phase-shifts.

Although the studies in [23], [31] extend earlier results on channel hardening to a wider scope of settings, they consider several simplifying assumptions which impact the validity of the final conclusions in practical scenarios. For instance, they consider the channel coefficients of reflecting elements to be *uncorrelated* and ignore the line-of-sight between the IRS and other terminals in the network. These assumptions are rather unrealistic, as in many use-cases of IRS-aided systems, the reflecting elements are assumed to be closely packed on the IRS, and the end-nodes are located in the line-of-sight of the IRS [38]. Moreover, [31] assumes the complete blockage of the direct path between the transmitter and the receiver which despite its validity in some use-cases, restricts the applicability of the results and makes the comparison between a large IRS-aided MIMO system and a massive MIMO system unfair. The asymptotic regime considered in [31] is further impractical, as we are often interested in IRS-aided MIMO settings with large IRS dimensions, but rather small arrays at the transmitter and receiver sides. In this work, we deviate from these simplifying assumptions and characterize the asymptotic channel hardening principle for a generic IRS-aided scenario.

C. Notation and Organization

Scalars, vectors and matrices are shown by non-bold, bold lower-case, and bold upper-case letters, respectively. The notation $\mathbf{H}^{\mathsf{H}}$ indicates the transposed conjugate of $\mathbf{H}$. An $N \times N$ identity matrix is denoted by $\mathbf{I}_N$ and $\mathbf{1}_N$ is an $N \times N$ matrix of all-ones. We use the notation $[\mathbf{H}]_{nm}$ to refer to the entry of $\mathbf{H}$ at the n-th row and m-th column. The function $\mathrm{Q}(x)$ is the standard Q-function, i.e.,

$$\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{+\infty} \mathrm{e}^{-\frac{u^2}{2}} \, \mathrm{d}u. \tag{1}$$

The mathematical expectation is denoted by $\mathcal{E}[\cdot]$, and the notation $\mathcal{CN}(\eta, \sigma^2)$ represents the complex Gaussian distribution with mean η and variance σ^2.

The rest of the manuscript is organized as follows: Section II describes the system model and formulates the problem. The main analytical results along with several numerical investigations are then presented in Section III. The dimensional trade-off between the transmitter and the IRS is investigated

in Section IV. Section V provides the detailed derivations of the main results. Finally, Section VI concludes the manuscript.

II. PROBLEM FORMULATION

Consider a MISO system in which a base station (BS) with an array-antenna of size M transmits data to a single-antenna user. An IRS with N reflecting elements is further employed to modify the propagation environment between the BS and the user: Each element of the IRS reflects its received signal after applying a phase-shift. The signal received by the user is then the superposition of two components: One that is received through the direct path between the BS and the user, and one that is reflected by the IRS.

A. System Model

As suggested by the literature [18], [32], we assume that the BS and the IRS are equipped with planar arrays. Namely, we assume the antenna-array at the BS and the array of reflecting elements on the IRS are rectangular planar arrays with M_{x} and N_{x} horizontal elements and M_{y} and N_{y} vertical elements, respectively, such that $M = M_{\mathrm{x}} M_{\mathrm{y}}$ and $N = N_{\mathrm{x}} N_{\mathrm{y}}$. The antenna elements at the BS are assumed to be distanced with ℓ_{x} and ℓ_{y} on the horizontal and vertical axes, respectively. The horizontal and vertical distances at the IRS are further denoted by d_{x} and d_{y}, respectively.

To model the direct path[2], we consider a classical scenario in which the line of sight (LoS) link is blocked via obstacles or mobility of the user. It is further assumed that ℓ_{x} and ℓ_{y} are set large enough, such that the spatial correlation can be ignored[3]. We therefore adopt the Rayleigh fading model in which the channel coefficient of the direct path between antenna element m at the BS and the user is modeled as

$$h_{\mathrm{d},m} = \sqrt{\alpha_{\mathrm{d}} A_M} \tilde{h}_{\mathrm{d},m}. \tag{2}$$

Here, α_{d} is a distance-dependent path-loss, A_M denotes the area of a single element on the planar antenna array at the BS, i.e., $A_M = \ell_{\mathrm{x}} \ell_{\mathrm{y}}$, and $\tilde{h}_{\mathrm{d},m}$ follows a circularly-symmetric complex Gaussian distribution with zero mean and unit variance, i.e., $\tilde{h}_{\mathrm{d},m} \sim \mathcal{CN}(0, 1)$. For sake of brevity, we define the direct channel vector as $\mathbf{h}_{\mathrm{d}} = [h_{\mathrm{d},1}, \cdots, h_{\mathrm{d},M}]^{\mathsf{T}}$.

Due to the flexibility in deploying the IRS, it is reasonable to assume that the link between the BS and the IRS is dominated by a LoS component. We denote the channel spanning from the BS to the IRS with $\mathbf{T} \in \mathbb{C}^{N \times M}$ where $t_{nm} = [\mathbf{T}]_{nm}$ denotes the channel coefficient between the m-th transmit antenna and the n-th IRS element. In particular, t_{mn} is given by

$$t_{nm} = \sqrt{\alpha_{\mathrm{s}} A_N} \bar{t}_{nm}, \tag{3}$$

where α_{s} represents the distance-dependent path-loss, A_N is the area of a single reflecting element, i.e., $A_N = d_{\mathrm{x}} d_{\mathrm{y}}$, and $\bar{t}_{nm}$ denotes the LoS components.

Unlike the direct and the BS-to-IRS links, the link between the user and the IRS has often both LoS and non-line of sight (NLoS) components. Moreover, the reflecting elements at the

---

[1]The difference of [23] to this work is illustrated shortly after.

[2]Note that the direct path is different from the line of sight.

[3]This is a rational assumption, as we do not assume M to be large.

IRS are often closely distanced, such that the given area of the IRS is filled with a large number of elements. As a result, the spatial correlation in this channel cannot be ignored. We hence use the Rician fading model with spatial correlation to model this channel: Let $h_{\mathrm{r},n}$ denote the coefficient of the channel between the n-th reflecting element and the user. This coefficient is modeled as

$$h_{\mathrm{r},n} = \sqrt{\alpha_{\mathrm{r}} A_N} \left(\sqrt{\frac{\kappa_{\mathrm{r}}}{\kappa_{\mathrm{r}}+1}} \bar{h}_{\mathrm{r},n} + \sqrt{\frac{1}{\kappa_{\mathrm{r}}+1}} \tilde{h}_{\mathrm{r},n} \right), \quad (4)$$

where α_{r} and κ_{r} are the distance-dependent path-loss and the Rician factor[1], respectively. The coefficient $\bar{h}_{\mathrm{r},n}$ represents the LoS, and $\tilde{h}_{\mathrm{r},n}$ models the small-scale fading process in the NLoS link. To capture the spatial correlation among the IRS elements, we consider a general covariance matrix $\mathbf{R} \in \mathbb{C}^{N \times N}$ for the vector $\tilde{\mathbf{h}}_{\mathrm{r}} = [\tilde{h}_{\mathrm{r},1}, \cdots, \tilde{h}_{\mathrm{r},N}]^{\mathsf{T}}$ which depends on d_{x} and d_{y}, distribution of scatterers, and the radiation pattern. For the case with isotropic scattering in front of the IRS, one can determine $\mathbf{R}$ explicitly from [32, Proposition 1]. Consequently, $\tilde{\mathbf{h}}_{\mathrm{r}}$ is modeled as a zero-mean complex Gaussian vector with covariance $\mathbf{R}$, i.e., $\tilde{\mathbf{h}}_{\mathrm{r}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$. Note that due to power normalization, we assume that $[\mathbf{R}]_{nn} = 1$ for $n \in [N]$.

Remark 1. *In general, the effective area of a single element on an array depends on the wave-length, and hence the given expressions for the array areas, i.e., $A_N = d_{\mathrm{x}} d_{\mathrm{y}}$ and $A_M = \ell_{\mathrm{x}} \ell_{\mathrm{y}}$, are not exact if the elements are distances longer than a half wave-length. Nevertheless, we assume the neighboring elements on the transmit array and the IRS to be distanced less than a half wave-length, i.e., $d_{\mathrm{x}}, d_{\mathrm{y}}, \ell_{\mathrm{x}}, \ell_{\mathrm{y}} \leq \lambda/2$.*

The LoS components are described by the array responses: Let λ be the wavelength and define the following operators

$$i_M(m) = (m-1) \mod M_{\mathrm{x}}, \quad (5a)$$
$$i_N(n) = (n-1) \mod N_{\mathrm{x}}, \quad (5b)$$

where $x \mod L$ determines x modulo L. Moreover, let

$$j_M(m) = \left\lfloor \frac{m-1}{M_{\mathrm{x}}} \right\rfloor, \quad (6a)$$
$$j_N(n) = \left\lfloor \frac{n-1}{N_{\mathrm{x}}} \right\rfloor. \quad (6b)$$

We now define an exponent function at azimuth angle φ and elevation angle θ for a given element m on the BS array and element n on the IRS array, respectively, as follows:

$$\Phi_m(\varphi, \theta) = i_M(m) \ell_{\mathrm{x}} \cos\theta \sin\varphi + j_M(m) \ell_{\mathrm{y}} \sin\theta, \quad (7a)$$
$$\Pi_n(\varphi, \theta) = i_N(n) d_{\mathrm{x}} \cos\theta \sin\varphi + j_N(n) d_{\mathrm{y}} \sin\theta, \quad (7b)$$

Consequently, the transmit and IRS array responses are given at (φ, θ) respectively by [32]

$$\mathbf{a}_M(\varphi, \theta) = \left[\mathrm{e}^{\frac{2\pi\mathrm{j}}{\lambda}\Phi_1(\varphi,\theta)}, \cdots, \mathrm{e}^{\frac{2\pi\mathrm{j}}{\lambda}\Phi_M(\varphi,\theta)} \right]^{\mathsf{T}}, \quad (8a)$$

$$\mathbf{a}_N(\varphi, \theta) = \left[\mathrm{e}^{\frac{2\pi\mathrm{j}}{\lambda}\Pi_1(\varphi,\theta)}, \cdots, \mathrm{e}^{\frac{2\pi\mathrm{j}}{\lambda}\Pi_N(\varphi,\theta)} \right]^{\mathsf{T}}. \quad (8b)$$

[1] In general, larger κ_{r} means that the channel is more deterministic.

Let $\bar{\mathbf{T}} \in \mathbb{C}^{N \times M}$ be the matrix of LoS channel coefficients between the BS and the IRS whose entry (n, m) is $\bar{t}_{nm}$. Define further the LoS channel vector between the IRS and the user as $\bar{\mathbf{h}}_{\mathrm{r}} = [\bar{h}_{\mathrm{r},1}, \cdots, \bar{h}_{\mathrm{r},N}]^{\mathsf{T}}$. We can hence write

$$\bar{\mathbf{T}} = \mathbf{a}_N(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1}) \mathbf{a}_M(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2})^{\mathsf{H}}, \quad (9a)$$
$$\bar{\mathbf{h}}_{\mathrm{r}} = \mathbf{a}_N(\varphi_{\mathrm{t}1}, \theta_{\mathrm{t}1}), \quad (9b)$$

where $(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1})$ is the angle-of-arrival (AoA) at the IRS, the pair $(\varphi_{\mathrm{t}1}, \theta_{\mathrm{t}1})$ denotes the angle-of-departure (AoD) from the IRS, and $(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2})$ is the AoD from the BS.

The signal received by the user is given by

$$y = \sum_{m=1}^{M} h_m x_m + z, \quad (10)$$

where z is zero-mean and unit-variance complex Gaussian noise, and x_m denotes the symbol sent by the m-th antenna element at the BS. The transmitted symbols satisfy

$$\sum_{m=1}^{M} \mathcal{E}\left[|x_m|^2 \right] \leq \rho, \quad (11)$$

for some total transmit power ρ. The coefficient h_m in (10) defines for the *end-to-end* effective channel between the m-th BS antenna and the user which is given by

$$h_m = h_{\mathrm{d},m} + \sum_{n=1}^{N} \mathrm{e}^{-\mathrm{j}\beta_n} h_{\mathrm{r},n} t_{nm}, \quad (12)$$

where β_n denotes the phase-shift with element n at the IRS.

B. Asymptotic Scaling of IRS

We intend to characterize the IRS-aided setting in a large-system limit in which the transmit-array size M remains fixed and the number of reflecting elements at the IRS grows large. In general, the growth in the number of reflecting elements can impact the physical dimension of the IRS in various forms. To illustrate this point, let A_{IRS} denote the total area of the IRS. Assuming the reflecting elements to be symmetrically inserted on the surface, we can write

$$A_{\mathrm{IRS}} = N A_N. \quad (13)$$

The scaling of A_{IRS} in terms of N is illustrated at best by considering the two most extreme scenarios.

- One extreme case is to assume that the number of IRS elements grows large and the distances between neighboring elements on the IRS are kept fixed and large enough, e.g., $d_{\mathrm{x}} = d_{\mathrm{y}} = \lambda/2$. In this case, A_N is fixed, and therefore the total area of the IRS grows linearly in N. Depending on the wave-length, this scaling can lead to unrealistic physical dimensions of IRS for large choices of N.
- Another extreme case is when we set the total area of the IRS fixed, and reduce the distance between the neighboring elements by the growth of N. In this case. A_{IRS} does not scale in N, and A_N shrinks reverse-linearly with N. In practice, however, the physical dimensions of a typical reflecting element cannot be set below a certain threshold. Hence, for a very large choice of N, assuming this extreme case can also be unrealistic.

In practice, with larger numbers of elements on the IRS, the distance among neighboring elements is reduced. This leads to a smaller effective area for a reflecting element and therewith to a higher spatial correlation. The distance can however be reduced down to a certain level. We can hence conclude that the scaling of the IRS area is in general neither linear nor fixed, but of some intermediate order; see some related discussions in [39], [40]. To take this point into account, we consider a basic scaling model for the IRS area. Namely, we assume that the area of each IRS element scales with N as

$$A_N = \frac{A_0}{N^q} \qquad (14)$$

for some constant A_0 and $0 \leq q \leq 1$. Consequently, the total area of the IRS scales with N as

$$A_{\mathrm{IRS}} = A_0 N^{1-q}. \qquad (15)$$

This basic scaling addresses the limiting scenarios in between the two extreme cases.

- $q = 0$ is the idealistic case which corresponds to the first extreme case, i.e., the case in which the distances between neighboring elements are kept fixed.
- $q = 1$ addresses the latter case in which the total area of the IRS is fixed.

By setting $0 < q < 1$, an intermediate scaling is considered in which the IRS surface grows sub-linearly by N.

It is worth mentioning that for $q \neq 1$, this scaling model implies that by sending $N \to \infty$, the area of IRS also grows asymptotically large. This observation can lead to this conclusion that with $q \neq 1$, the far-field derivation for the IRS array response is no longer valid in the large-system limit. To avoid such inconsistency, let $d_{\min}$ denotes the minimum distance between two terminals in the network, i.e., the minimum of the BS-to-user, BS-to-IRS and IRS-to-user distances. We assume that $d_{\min}$ is bounded uniformly from below as

$$d_{\min} > D_0 N^{\gamma/2}, \qquad (16)$$

for some D_0 and $\gamma > 1 - q$. By considering this assumption, we guarantee that the far-field derivations are valid through the asymptotic analyses. From the implementational point of view, this is a realistic assumption, as the distances are often in orders of tens or hundreds of meters while the IRS dimensions are often a finite multiple of the wave-length.

C. Input-Output Mutual Information

Let $\mathbf{h} = [h_1, \cdots, h_M]^{\mathsf{T}}$ denote the end-to-end channel vector. Considering Gaussian signaling, the mutual information between the input and output of the end-to-end channel, for a given realization of $\mathbf{h}$, is given by

$$\mathcal{I}\left(\mathbf{h}, \rho \mathbf{Q}\right) = \log_2\left(1 + \rho \mathbf{h}^{\mathsf{H}} \mathbf{Q} \mathbf{h}\right), \qquad (17)$$

where $\mathbf{Q} \in \mathbb{C}^{M \times M}$ is the transmit covariance matrix. From the power constraint, we know that $\mathbf{Q}$ satisfies $\mathrm{tr}\{\mathbf{Q}\} = 1$.

In this channel, the optimal covariance matrix which maximizes the mutual information is given by maximum ratio transmission (MRT) [41], [42]. We hence focus on the optimal

case which determines the capacity of the end-to-end channel. To this end, we define the maximum mutual information[1] as

$$\mathcal{C} = \log_2\left(1 + \rho M \Gamma\right), \qquad (18)$$

which determines the end-to-end channel capacity. Here, Γ is given by

$$\Gamma = \frac{\|\mathbf{h}\|^2}{M}. \qquad (19)$$

In the sequel, we refer to Γ as the *end-to-end SNR gain* per transmit antenna. Due to fading, Γ and $\mathcal{C}$ are random. Our main goal is to characterize the statistics of these random variables analytically, when N is asymptotically large.

III. ASYMPTOTIC CHANNEL HARDENING

Intuitively, channel hardening in massive MIMO systems refers to the following phenomenon: A fading MIMO channel behaves almost deterministically, as the number of antennas at one end grows large[2]. This fundamental property is characterized in the seminal work [1] in which a tight approximation for the distribution of the input-output mutual information is determined. The result shows that the variance of the mutual information shrinks rapidly while its mean grows large.

Invoking the literature, it can be straightforwardly shown that the end-to-end channel $\mathbf{h}$ hardens, as the number of transmit antennas M grows unboundedly large. We are however interested in a different asymptotic regime; namely, a scenario with *finite* transmit antennas but unboundedly *large number of IRS elements*, i.e., fixed M and $N \to \infty$. Our main result shows that in this alternative asymptotic regime, the channel still hardens, under a very mild constraint on the correlation among the IRS elements.

Proposition 1. *Let the area of the IRS scale with N as $A_{\mathrm{IRS}} = A_0 N^{1-q}$ for some fixed A_0 and $0 \leq q < 1$, and the phase-shifts be set to*

$$\beta_n^\star = \frac{2\pi}{\lambda}\left(\Pi_n\left(\varphi_{\mathrm{r1}}, \theta_{\mathrm{r1}}\right) + \Pi_n\left(\varphi_{\mathrm{t1}}, \theta_{\mathrm{t1}}\right)\right). \qquad (20)$$

Assume the maximum eigenvalue of the IRS covariance matrix $\mathbf{R}$, denoted by $\lambda_{\max}$, grows with N sub-linearly, i.e.,

$$\lim_{N \to \infty} \frac{\lambda_{\max}}{N} = 0, \qquad (21)$$

and satisfies

$$\lim_{N \to \infty} \frac{1}{\lambda_{\max} A_{\mathrm{IRS}}} = 0. \qquad (22)$$

Then, the maximum mutual information $\mathcal{C}$ is well approximated by a real Gaussian random variable whose mean and standard deviation are given by

$$\mu_{\mathcal{C}} = \log_2\left(1 + \rho M \mu\right) \qquad (23\mathrm{a})$$

$$\sigma_{\mathcal{C}} = \frac{\rho M \log_2 \mathrm{e}}{1 + \rho M \mu} \sqrt{\omega\eta + \eta + \frac{M-1}{M}\alpha_{\mathrm{d}} A_M} \qquad (23\mathrm{b})$$

---

[1]We show it by $\mathcal{C}$, as it determines the end-to-end *capacity*.
[2]In general, it is enough to have a large antenna array at one end.

respectively, for

$$\mu = \alpha_{\mathrm{d}} A_M + \kappa_{\mathrm{r}} \bar{\alpha}_N N^2 + \bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}}, \qquad (24\mathrm{a})$$

$$\eta = \frac{\alpha_{\mathrm{d}} A_M}{M} + \bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}}, \qquad (24\mathrm{b})$$

$$\omega = 2\kappa_{\mathrm{r}} \bar{\alpha}_N N^2 + \bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}}, \qquad (24\mathrm{c})$$

and $\bar{\alpha}_N$ being defined as

$$\bar{\alpha}_N = \frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} A_N^2. \qquad (25)$$

Proof. The proof is given in Section V-C. $\qquad\square$

Proposition 1 considers three main constraints on the system setting; namely, it restricts the scaling order of the IRS area, assumes sub-linearly growing $\lambda_{\max}$ and the constraint in (22). In this respect, it is worth mentioning few remarks.

- Proposition 1 excludes the limiting case of $q = 1$, for analytical rigor. Nevertheless, as we see in the forthcoming section, the analytical expressions are still rather accurate for $q = 1$.
- Constraining the growth rate of the maximum eigenvalue $\lambda_{\max}$ to be sub-linear can be interpreted as follows: $\lambda_{\max}$ is uniformly bounded from above as $\lambda_{\max} \leq aN^u$ for some real scalars a and $0 \leq u < 1$. This is in general not a strong constraint, as we can write

$$\lambda_{\max} \leq \mathrm{tr}\{\mathbf{R}\} = N. \qquad (26)$$

As we illustrate later on, the exponent u is mutually coupled with q. This is seen by considering the two extreme cases of q. For $q = 0$, it is feasible to have $\mathbf{R} = \mathbf{I}_N$, and hence $\lambda_{\max} = 1$ meaning that $u = 0$. The other extreme case is $q \to 1$ in which $\mathbf{R}$ tends to a rank-one matrix. In this case $\lambda_{\max} = N$, and hence the uniform upper-bound is valid with[1] $a = u = 1$. For less rank-deficient covariance matrices, u can be something in $[0, 1]$. Similar to the constraint on the physical dimension of the IRS, we show later that despite assuming $0 \leq u < 1$, Proposition 1 is still valid for the extreme case of $u = 1$.
- Considering the scaling order of the IRS area A_{IRS} and maximum eigenvalue $\lambda_{\max}$, the constraint in (22) restricts u and q to satisfy $u \geq q$. One may find this constraint intuitively valid following from the above discussions on the coupling of u and q. We later on show that this is in fact the case for the covariance matrix derived for the standard Rayleigh model in [32].

A. Numerical Investigations

Before starting with derivations, let us confirm the accuracy of Proposition 1 for practical system dimensions through some numerical experiments. To this end, we consider a basic scenario in which the BS is equipped with a 2×2 planar array and the IRS surface consists of $N = 256$ reflecting elements which are aligned on a rectangular surface with $N_{\mathrm{x}} = 8$ horizontal elements and $N_{\mathrm{y}} = 32$ vertical elements. The elements at both BS and IRS are distanced with $\ell_{\mathrm{x}} = d_{\mathrm{x}} = \ell_{\mathrm{y}} = d_{\mathrm{y}} = \lambda/2$.

[1]Similar to q, we exclude the linear case, i.e., $u = 1$ for analytical rigor.
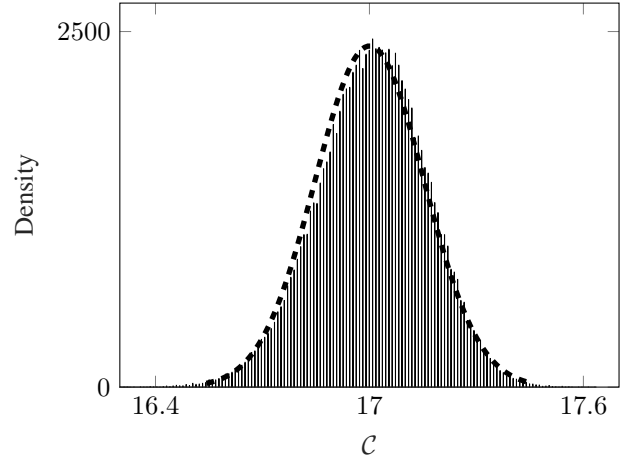

Fig. 1: Histogram of the mutual information and the fitted Gaussian distribution for the covariance matrix $\mathbf{R}$ given in (27).

For sake of brevity, we set $\alpha_{\mathrm{d}} A_M = \alpha_{\mathrm{r}} A_N = \alpha_{\mathrm{s}} A_N = 1$, and $\log \kappa_{\mathrm{r}} = 0$ dB. For the covariance matrix $\mathbf{R}$, we consider a Rayleigh fading model and invoke [32, Proposition 1]. The entry (n, n') of $\mathbf{R}$ is hence given by

$$[\mathbf{R}]_{nn'} = \mathrm{sinc}\left(\frac{2}{\lambda}\sqrt{D_{\mathrm{x}}^2 + D_{\mathrm{y}}^2}\right) \qquad (27)$$

with D_{x} and D_{y} being

$$D_{\mathrm{x}} = d_{\mathrm{x}} \left[i_N(n) - i_N(n')\right], \qquad (28\mathrm{a})$$

$$D_{\mathrm{y}} = d_{\mathrm{y}} \left[j_N(n) - j_N(n')\right]. \qquad (28\mathrm{b})$$

The AoA and AoDs are further set to $(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1}) = (\pi/6, \pi/3)$, $(\varphi_{\mathrm{t}1}, \theta_{\mathrm{t}1}) = (\pi/8, 2\pi/3)$ and $(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2}) = (\pi/7, \pi/5)$. The power constraint is set to $\rho = 1$.

For this setting, we collect 10^5 realizations of the channel and determine the input-output mutual information for every realization. The histogram of the collected samples is shown in Fig. 1. As the figure shows, the histogram closely matches the Gaussian distribution. We now plot the analytical distribution given by Proposition 1 and compare it to the Gaussian distribution fitted to the histogram. The result is shown in Fig. 2. As the figure demonstrates, the analytical result of Proposition 1 matches almost perfectly to the empirical density.

We now conduct a new experiment. We consider the same setting and let the IRS size to vary as $N = N_{\mathrm{x}}^2$ while changing from $N_{\mathrm{x}} = 8$ to $N_{\mathrm{x}} = 36$, gradually. Note that in this setting, the area of the IRS grows linearly[2] in N, i.e., $t = 0$, and hence the limit in (22) is satisfied. For each choice of N, we collect 500 samples, determine numerically the mean and variance of $\mathcal{C}$ and plot it against N in Figs. 3 and 4.

The numerical results in Figs. 3 and 4 are compared with the closed-form expressions given in Proposition 1. As the figures show, the analytical expressions closely track the numerical results. The figures further demonstrate the hardening of the channel in terms of the IRS size. As N grows large, the mean mutual information grow large, while the variance drops.

[2]Such scaling is not realistic for very large IRSs. We discuss more realistic scaling scenarios in the forthcoming parts of this section.
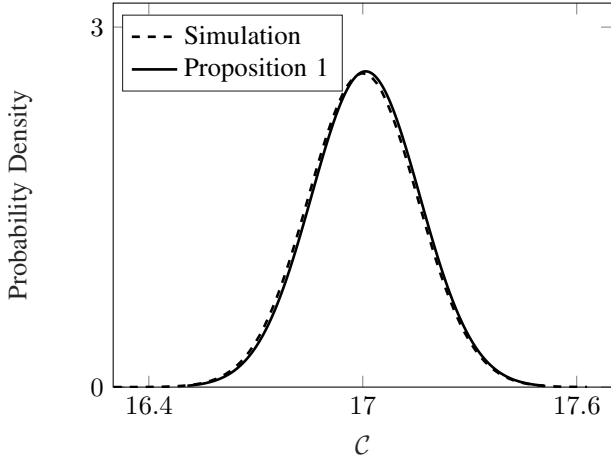
Fig. 2: Probability density of the maximum mutual information and the Gaussian distribution fitted to the numerical simulations.
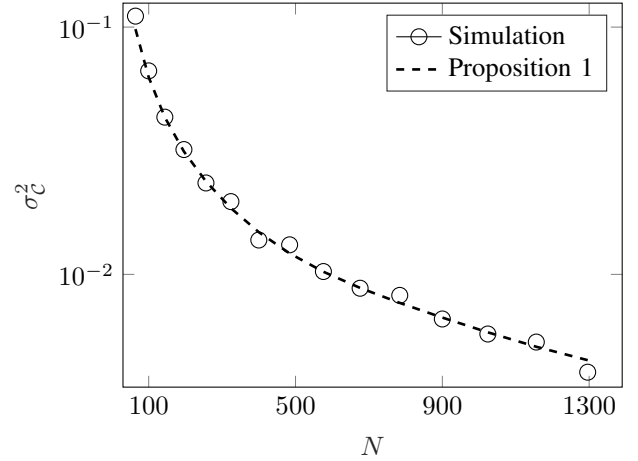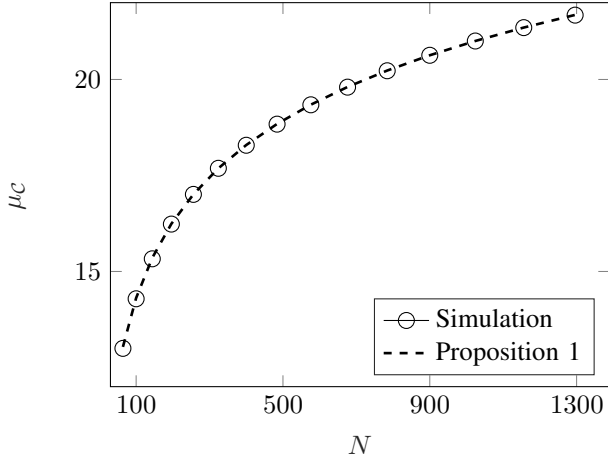

Fig. 3: Numerical simulations and the analytic expression for the mean mutual information against N.

B. Channel Hardening Order

The numerical experiments in the previous section consider the case of $q = 0$, i.e., when the IRS area grows linearly in N. Although this can be considered feasible for a small or moderate number of reflecting elements, it is not a realistic scaling for asymptotically large IRSs. In fact, in such scenarios, the total area of the IRS is limited, and hence, the area grows sub-linearly, i.e., $q > 0$. We address this point by investigating the speed of channel hardening in this section. We are mainly interested to find out how fast the channel hardens with respect to N for a given scaling of the IRS area. In general, the speed depends on the correlation among the reflecting elements, and therefore the physical dimensions of the IRS. This argument is analytically characterized in Proposition 2.

Proposition 2. *Let the sub-linearly growing maximum eigenvalue of* $\mathbf{R}$ *be uniformly bounded from above as*

$$\lambda_{\max} \leq a N^u, \tag{29}$$

for some non-negative real a *and* $0 \leq u < 1$*. Then, there exists an integer* N_0*, such that for* $N \geq N_0$ *the mean and variance*

Fig. 4: Numerical simulations and the analytic expression for the variance of the mutual information against N.

of the maximum mutual information is bounded uniformly as

$$\mu_{\mathcal{C}} \geq b + (1 - q) \log_2 N \tag{30a}$$

$$\sigma_{\mathcal{C}}^2 \leq \frac{c}{N^{1-u}}, \tag{30b}$$

for some non-negative real scalars b *and* c*.*

Proof. We start the proof by noting that

$$\bar{\alpha}_N = \frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} A_N^2 \tag{31a}$$

$$= \frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} \frac{A_{\mathrm{IRS}}^2}{N^2} = \frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} \frac{A_0^2}{N^{2q}}. \tag{31b}$$

From Proposition 1, since $\mu \geq \bar{\alpha}_N \kappa_{\mathrm{r}} N^2$, we have

$$\mu_{\mathcal{C}} \geq \log_2 \left(\rho M \kappa_{\mathrm{r}} \bar{\alpha}_N N^2 \right). \tag{32}$$

Using (31b), we can hence conclude that

$$\mu_{\mathcal{C}} = \log_2 \left(\frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} \rho M \kappa_{\mathrm{r}} A_0^2 \right) + \log_2 N^{2-2q}. \tag{33}$$

Noting that the first term on the right hand side does not grow in N, we conclude that

$$\mu_{\mathcal{C}} \geq b + (1 - q) \log_2 N, \tag{34}$$

by setting b to

$$b = \log_2 \left(\frac{\alpha_{\mathrm{r}} \alpha_{\mathrm{s}}}{1 + \kappa_{\mathrm{r}}} \rho M \kappa_{\mathrm{r}} A_0^2 \right). \tag{35}$$

To further bound the variance, we note that $\left\| \bar{\mathbf{h}}_{\mathrm{r}} \right\|^2 = N$. As a result, $\bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}} / N$ determines the *Rayleigh quotient* of $\mathbf{R}$ at $\bar{\mathbf{h}}_{\mathrm{r}}$ which is bounded from above by the maximum eigenvalue of $\mathbf{R}$ [43]. We hence have

$$\bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}} \leq \bar{\alpha}_N \lambda_{\max} N. \tag{36}$$

Let $\lambda_{\max}$ be bounded uniformly from above by $a N^u$ for some $0 \leq u < 1$. This concludes that μ, η and ω in Proposition 1 are uniformly bounded from above respectively by

$$\mu^{\mathrm{U}}(N) = \mathcal{O}\left(N^{2-2q}\right) \tag{37a}$$

$$\eta^{\mathrm{U}}(N) = \mathcal{O}\left(N^{1+u-2q}\right) \tag{37b}$$

$$\omega^{\mathrm{U}}(N) = \mathcal{O}\left(N^{2-2q}\right). \tag{37c}$$

We now start with bounding $\sigma_{\mathcal{C}}$. Since $1 + \rho M \mu > \rho M \mu$, we can initially bound the variance as

$$\sigma_{\mathcal{C}} < \frac{\log_2 e}{\mu} \sqrt{\omega \eta + \eta + \frac{M-1}{M} \alpha_d A_M} = \Xi_N \sqrt{\frac{\eta}{\mu}} \quad (38)$$

where Ξ_N is given by

$$\Xi_N = \sqrt{\frac{1}{\mu \eta} \left(\omega \eta + \eta + \frac{M-1}{M} \alpha_d A_M \right) \log_2 e} \quad (39)$$

The term Ξ_N is uniformly bounded from above by a constant. We can hence conclude that there exists an integer N_1 and a non-negative real c_0, such that for $N \geq N_1$, we have

$$\sigma_{\mathcal{C}} \leq c_0 \sqrt{\frac{\eta}{\mu}} \quad (40)$$

Noting that $\mu \geq \kappa_r \bar{\alpha}_N N^2$, we can further bound the above upper-bound as

$$\sigma_{\mathcal{C}} \leq c_1 \frac{\sqrt{\eta}}{N^{1-q}} \quad (41)$$

where c_1 being defined as

$$c_1 = \sqrt{\frac{c_0 (1 + \kappa_r)}{\alpha_r \alpha_s \kappa_r}}. \quad (42)$$

Since η is uniformly bounded with $\eta^U(N)$, we conclude that there exists an integer $N_2 \geq N_1$ and a non-negative real c_2, such that for $N \geq N_2$, we have

$$\frac{\eta}{N^{2-2q}} \leq \frac{c_2}{N^{1-u}}. \quad (43)$$

The proof is finally completed by setting $c = c_1^2 c_2$ and N_0 to be $N_0 = N_2$. $\square$

Proposition 2 formulates an intuitive behavior. The higher the correlation among the IRS elements is, i.e., larger u, the slower the end-to-end channel hardens. It further indicates that the fastest hardening is achieved by $u = q = 0$. This corresponds to reflecting arrays whose number of *strongly* correlated elements does not scale with N while the area of the IRS grows linearly in N. From implementation viewpoint, this means that the neighboring reflecting elements on the IRS are well-distanced, and the physical dimensions of the IRS grow with the number of reflecting elements, such that the distance between two neighboring elements remains constant.

We now validate Proposition 2 through a numerical experiment. We consider the covariance matrix derived in [32, Proposition 1] for a planar IRS whose elements are spaced with $\lambda/2$ in both directions, i.e., $d_x = d_y = \lambda/2$. This means that for this array $q = 0$. The covariance matrix in this case is specified via (27). We consider a square array, i.e., $N_x = N_y$ and let N grow gradually from $N = 64$ to $N = 1296$. For this sequence of covariance matrices, λ_{max} is plotted against N in Fig. 5. We further use the curve fitting toolbox of MATLAB, i.e., `cftool` [44], to fit the collected data to a curve of form $\lambda_{max} = aN^u$. The result is shown with a dashed line in the figure for which $a = 0.83$ and $u = 0.25$.

Given the results in Fig. 5, Proposition 2 suggests that the variance of the mutual information in this case is uniformly
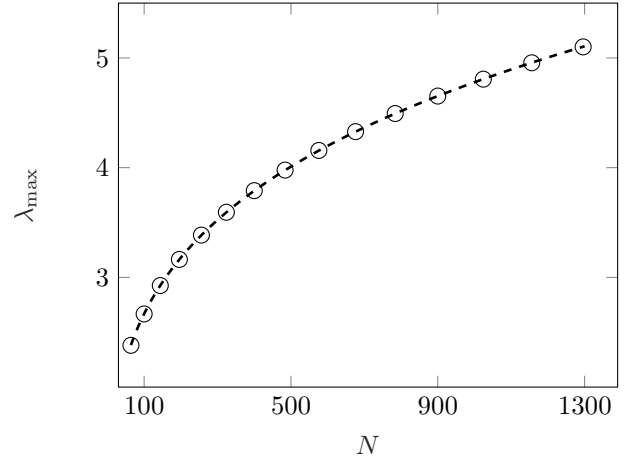

Fig. 5: λ_{max} against N for the classic IRS covariance matrix derived in [32, Proposition 1]. The dashed line shows the fitted curve which reads $\lambda_{max} \approx 0.83 \sqrt[4]{N}$, i.e., $u = 0.25$.
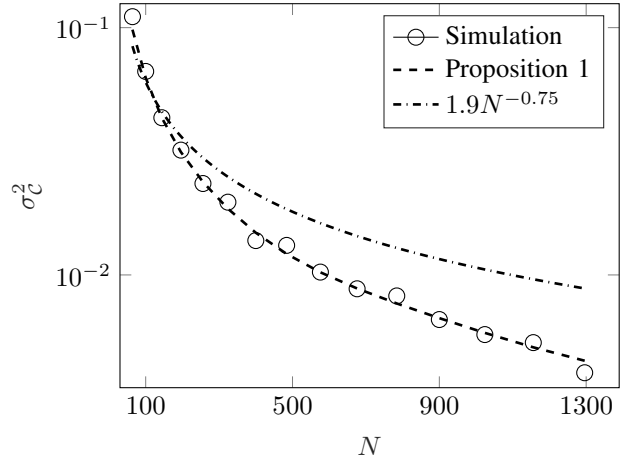

Fig. 6: Comparing the channel hardening speed to the uniform upper bound suggested by Proposition 2.

bounded from above by $cN^{-0.75}$ for some c. This is shown in Fig. 6, where we compare $\sigma_{\mathcal{C}}^2$ against the upper bound with $c = 1.9$. The numerical results show consistency with Proposition 2. Interestingly, the suggested upper bound gives a *pessimistic* approximation of the hardening speed. In fact the true variance, drops much faster than the order of the upper bound.

C. Scaling of IRS Area and Covariance Matrix

As indicated earlier, u and q are mutually coupled. This is easily seen by sending $q \to 1$. In this case, the total area of the IRS is constant. This means that in the limit $N \to \infty$, the covariance matrix converges to a rank-one matrix and hence

$$\lambda_{max} = \text{tr}\{\mathbf{R}\} = N, \quad (44)$$

or equivalently $u \to 1$; see also [32]. To understand this mutual coupling analytically and find out how realistic the condition (22) is, i.e., $u \geq q$, we conduct a numerical experiment. For a given q, we repeat the previous experiment. This means that we consider square IRSs of size N, i.e., $N_x = N_y = \sqrt{N}$, and let N grow gradually from $N = 64$ to $N = 1296$. The
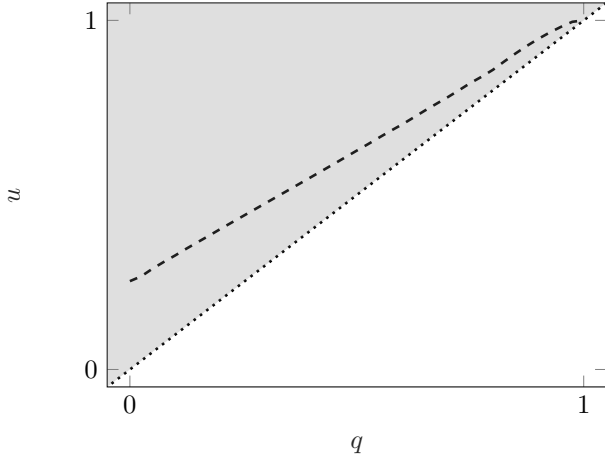
Fig. 7: Eigenvalue order exponent u against the area order exponent q. The gray area is the feasible region $u \geq q$ where (22) is satisfied.

Fig. 8: Histogram and the fitted Gaussian density of $\mathcal{C}$ for the extreme case of $\mathbf{R} = \mathbf{1}_N$. The scaled version of the analytical density function given by Proposition 1 is further plotted with a solid line.

distance between two neighboring elements on the IRS in this case is

$$d_{\mathrm{x}} = d_{\mathrm{y}} = \frac{\lambda}{2N^{q/2}}. \tag{45}$$

Invoking the least-squares method, we numerically find u such that $\lambda_{\max} \leq aN^u$ for some real a. We then change q from $q = 0$ to $q = 1$ and plot u against it. The result is shown in Fig. 7 along with the region $u \geq q$ which corresponds to the constraint given by (22).

As the figure suggests, for the covariance matrix, derived for the Rayleigh fading model, $u \geq q$. This is intuitive, since in both extreme cases of $q = 0$ and $q \to 1$, this constraint is satisfied. The result further show a linear growth of u in q. This observation further suggests that Propositions 1 and 2 are further valid approximations in the limits $q \to 1$ and $u \to 1$. We confirm this intuition through numerical analysis of this limiting case in the sequel.

D. The Extreme Case of Rank-One Covariance

As mentioned, the sub-linearity constraint in Propositions 1 and 2 is not a strong constraint. In fact, noting that $\lambda_{\max} \leq N$, one can conclude that the sub-linearity constraint is satisfied in all settings, but the *extreme* case with a *finite-rank* covariance matrix whose rank does not grow with N, and therefore $\lambda_{\max}$ scales linearly in N. From the mutual coupling between u and q, in this case we further have $q = 1$ which corresponds to an IRS whose total area is kept fixed.

Inspired by our observation in Section III-B, we conjecture that Propositions 1 and 2 give accurate characterizations, even in an extreme case with a linearly growing $\lambda_{\max}$. To confirm this conjecture, we run simulations for the same setting considered in Section III-A while replacing the covariance matrix with $\mathbf{R} = \mathbf{1}_N$, i.e., the matrix of all ones, and correspondingly setting $q = 1$, i.e., assuming
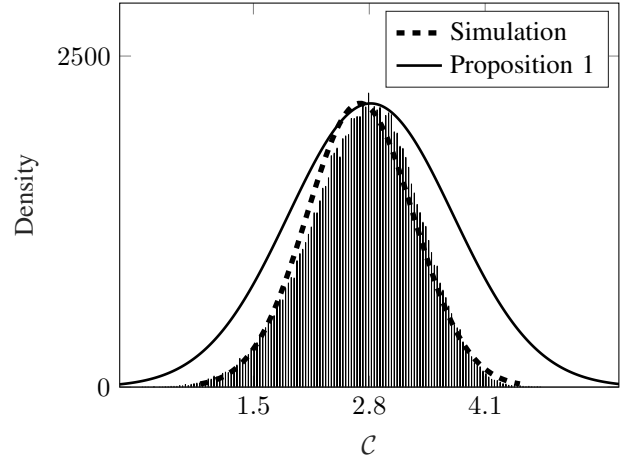
$$A_N = \frac{A_0}{N}. \tag{46}$$

for some real A_0. To achieve this scaling, we set the distance between two neighboring IRS elements to be

$$d_{\mathrm{x}} = d_{\mathrm{y}} = \frac{\lambda}{2\sqrt{N}} \tag{47}$$

which means that $A_0 = \lambda^2/4$. We further set $\alpha_{\mathrm{r}} = \alpha_{\mathrm{s}} = 1/A_0$ to make the comparison to the first experiment in Section III-A fair. The remaining parameters are the same as those considered in Section III-A.

The matrix $\mathbf{R} = \mathbf{1}_N$ is a rank one matrix whose maximum eigenvalue reads $\lambda_{\max} = N$, and thus scales linearly with N. One can observe this case, as an extreme case of the covariance matrix in (27), in which the distance between the most outer two elements of the IRS is still significantly smaller than the wavelength[1].

Fig. 8 shows the numerically-evaluated histogram of $\mathcal{C}$, as well as the properly scaled version of the fitted Gaussian probability density and the distribution suggested by Proposition 1. The histogram is evaluated for 10^5 channel realizations. As the figure shows, the histogram in this case is rather loosely approximated by the distribution given in Proposition 1, and compared to the almost-perfect match in Figs. 1 and 2, it shows some approximation error. This approximation error follows from the lower speed of channel hardening in this case. In fact, from Proposition 2, we know that by sending $q, u \to 1$, the mean and variance of $\mathcal{C}$ scale with $\mathcal{O}(1)$ in N, and hence, the approximation in Proposition 1 becomes inaccurate[2].

Although the approximation by Proposition 1 is inaccurate, the histogram in Fig. 8 shows a rather small standard deviation for $\mathcal{C}$. This is an interesting observation confirming our earlier finding in Section III-B which indicates that the upper bound given by Proposition 2 is rather *pessimistic*. This finding can be analytically supported as follows: The sub-linearity constraint in Propositions 1 and 2 comes from utilizing the inequality $\bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}} \leq \lambda_{\max} N$ to bound the variance $\sigma_{\mathcal{C}}^2$. Nevertheless, for

---

[1] In other words, an IRS whose physical dimensions are significantly smaller than the wavelength.

[2] This is the main reason that Proposition 1 excludes the case $q = u = 1$.
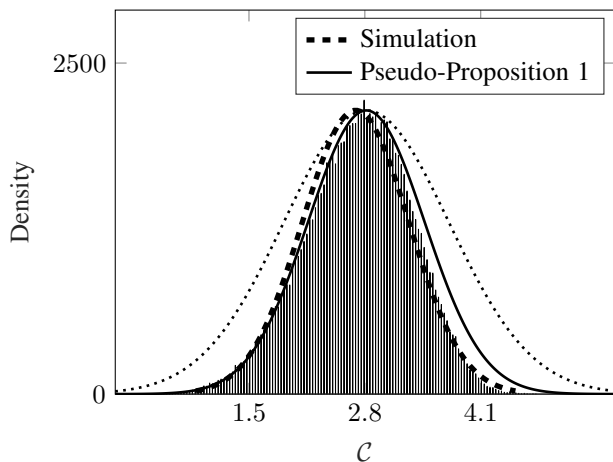
Fig. 9: Histogram and the fitted Gaussian density of $\mathcal{C}$ for the extreme case of $\mathbf{R} = \mathbf{1}_N$. The scaled version of the density functions given by Pseudo-Proposition 1 and Proposition 1 are plotted with solid and dotted lines, respectively.

rank-deficient covariance matrices, this bound is often asymptotically very loose. This means that despite the inaccuracy of Proposition 1 for exactly linear scaling, i.e., $q = u = 1$, the channel hardening still occurs in this case.

We now invoke the above heuristic conclusion to further modify Proposition 1 for linearly scaling A_{IRS} and λ_{max}. In fact, assuming that $\mathcal{C}$ is still Gaussian in this case, we can conclude that the approximation error in this case comes from the inaccuracy of the variance in Proposition 1 for cases with linear scaling. The following pseudo-proposition gives a modified approximation for the distribution of $\mathcal{C}$ which is also accurate for linearly scaling scenarios. Interestingly, by assuming sub-linear scaling, the approximation reduces to the limiting distribution given by Proposition 1.

Pseudo-Proposition 1. *For linearly scaling A_{IRS} and λ_{max}, i.e., cases with $q = u = 1$, the maximum mutual information $\mathcal{C}$ is well approximated by a real Gaussian random variable whose mean is $\mu_{\mathcal{C}}$ and whose variance is*

$$\hat{\sigma}_{\mathcal{C}}^2 = \frac{\kappa_{\text{r}}}{\kappa_{\text{r}} + \vartheta} \sigma_{\mathcal{C}}^2, \tag{48}$$

where $\mu_{\mathcal{C}}$ and $\sigma_{\mathcal{C}}^2$ are given in Proposition 1, and

$$\vartheta = \frac{\alpha_{\text{d}} A_M}{2 \bar{\alpha}_N N^2} \tag{49}$$

Proof. The proof is given in Section V-D. $\qquad\square$

For the numerical experiment of Fig. 8, we further plot the approximation by Pseudo-Proposition 1 in Fig. 9. As the figure demonstrates, the proposed approximated density tracks the empirically-evaluated histogram more closely[1] compared with the limiting result given by Proposition 1.

IV. IRS-TRANSMITTER DIMENSIONAL TRADE-OFF

The analytical results of this study enable us to address various design challenges in IRS-aided systems. In this section,

[1] The interested reader can check that for $u, q < 1$, we have $\hat{\sigma}_{\mathcal{C}}^2 \to \sigma_{\mathcal{C}}^2$, as we set $N \to \infty$.

we focus on a particular application: We employ the results to investigate the dimensional trade-off between the BS and the IRS. More precisely, we try to answer this fundamental question: How does the transmit array dimension, i.e., M, change, when we employ an IRS to enhance the communication link?

A. Dimensional Trade-Off for Ergodic Capacity

We start our investigations by considering cases in which the NLoS links experience a fast fading process. This means that the channel varies from a short block of symbol transmit intervals to another. In this case, the performance is best described by the *ergodic capacity*, defined as

$$\bar{\mathcal{C}} = \mathcal{E}\left[\mathcal{C}\right]. \tag{50}$$

For a target ergodic capacity, the BS needs to be equipped with a certain number of transmit antennas, i.e., M. Intuitively, this number of required antennas is expected to reduce, as we increase the IRS dimension. This draw a dimensional trade-off between N and M: With a larger IRS, the minimum number of required BS antennas to achieve a given target performance decreases. Proposition 1 enables us to quantitatively formulate this dimensional trade-off. To this end, we first use Proposition 1 and derive the ergodic capacity for a given N and M in a closed form as follows:

$$\bar{\mathcal{C}} = \log_2\left(1 + \rho M \left[\alpha_{\text{d}} A_M + \kappa_{\text{r}} \bar{\alpha}_N N^2 + \bar{\alpha}_N \bar{\mathbf{h}}_{\text{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\text{r}}\right]\right). \tag{51}$$

It is hence readily concluded that for a given N and the target ergodic capacity $\bar{\mathcal{C}}$, M needs to satisfy

$$M \geq \frac{2^{\bar{\mathcal{C}}} - 1}{\rho \left(\alpha_{\text{d}} A_M + \kappa_{\text{r}} \bar{\alpha}_N N^2 + \bar{\alpha}_N \bar{\mathbf{h}}_{\text{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\text{r}}\right)}. \tag{52}$$

The inequality in (52) formulates the trade-off between N and M: By increasing the physical dimensions of the IRS, a smaller array is required at the BS. It further specifies the speed of this drop, i.e., M drops proportional to N^{2q-2}.

We now conduct some experiments to investigate this trade-off numerically. To this end, a setting is considered in which an IRS is distanced from the BS with $D_{\text{s}} = 25$ m. The receiver is further located at a random point at which its distances from the BS and the IRS are $D_{\text{d}} = 20$ m and $D_{\text{d}} = 15$ m, respectively. We further set the AoA and AoDs to be the same as those considered in Section III-A, and assume that $\rho = 1$. The large-scale fading coefficients are generated as

$$\alpha_i = \frac{\alpha_{\text{ref}}}{D_i^{\varepsilon_i}} \tag{53}$$

for $i \in \{\text{s}, \text{d}, \text{r}\}$, where ε_i is the path-loss exponent of link i and α_{ref} is the reference path-loss. For numerical simulations, we set $\log \alpha_{\text{ref}} = 10$ dB, $\varepsilon_{\text{d}} = 3.5$ and $\varepsilon_{\text{s}} = \varepsilon_{\text{r}} = 2.3$.

The IRS is considered to be a square planar array, i.e., $N_{\text{x}} = N_{\text{y}}$, whose elements are distanced with a half wavelength. This means that $q = 0$. The covariance matrix of the NLoS link is further set to be the Rayleigh covariance matrix, i.e., (27). We assume that the phase-shifts of the reflecting elements are set according to Proposition 1; hence, the configuration of the transmit array is not required to be taken into account.

To investigate the dimensional trade-off between M and N, we let the ergodic capacity to be set to a given target value $\bar{\mathcal{C}}$.
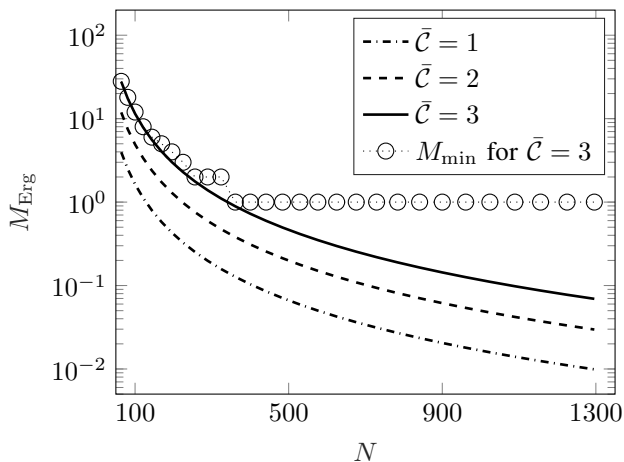
Fig. 10: IRS-transmitter dimensional trade-off for target ergodic capacity $\bar{\mathcal{C}}$: The minimum required number of transmit antennas $M_{\min}$ is determined by quantizing M_{Erg} to the next larger integer.

Fig. 11: IRS-transmitter dimensional trade-off for various target outages: Here, the target rate is set to $\mathcal{R} = 3$.

The number of reflecting elements is then increased gradually starting from $N = 64$ and ending at $N = 1296$. For each N, we determine the *real-valued* lower bound on the minimum required M, denoted by M_{Erg}, from (52), i.e.,

$$M_{\mathrm{Erg}} = \frac{2^{\bar{\mathcal{C}}} - 1}{\rho \left(\alpha_{\mathrm{d}} A_M + \kappa_{\mathrm{r}} \bar{\alpha}_N N^2 + \bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}} \right)}. \quad (54)$$

Fig. 10 plots M_{Erg} against the IRS size N, for three various choices of $\bar{\mathcal{C}}$. The minimum required BS array size, denoted by $M_{\min}$, is then determined by quantizing M_{Erg} to the next larger integer. As the figure shows, for rather large choices of N, the target ergodic capacity with only one antenna at the transmitter. Our further numerical investigations confirm the consistency of the result with simulations.

B. Dimensional Trade-Off for Outage Probability

With a slow fading process, the channel remains fixed within a long sequence of symbol intervals. The appropriate metric for performance evaluation in this case is therefore the outage probability which is defined for a given target rate $\mathcal{R}$ as [45]

$$\mathcal{P}_{\mathrm{out}} (\mathcal{R}) = \Pr \{ \mathcal{C} < \mathcal{R} \} . \quad (55)$$

The dimensional trade-off can be also studied for this metric using the analytical results of Section III. From Proposition 1, we can determine the outage probability as

$$\mathcal{P}_{\mathrm{out}} (\mathcal{R}) = Q \left(\frac{\mu_{\mathcal{C}} - \mathcal{R}}{\sigma_{\mathcal{C}}} \right), \quad (56)$$

with $Q (\cdot)$ being the standard Q-function. To guarantee achieving $\mathcal{P}_{\mathrm{out}} (\mathcal{R}) \leq p_{\mathrm{out}}$ for a given target rate $\mathcal{R}$, we need to have

$$\frac{\mu_{\mathcal{C}} - \mathcal{R}}{\sigma_{\mathcal{C}}} \geq Q^{-1} (p_{\mathrm{out}}), \quad (57)$$

where $Q^{-1} (\cdot)$ is the inverse of the Q-function with respect to composition, i.e., $Q^{-1} (Q^{-1} (p_{\mathrm{out}})) = p_{\mathrm{out}}$.

For a given reflecting array and rate $\mathcal{R}$, it is readily shown, after few lines of calculations, that the target outage is achieved
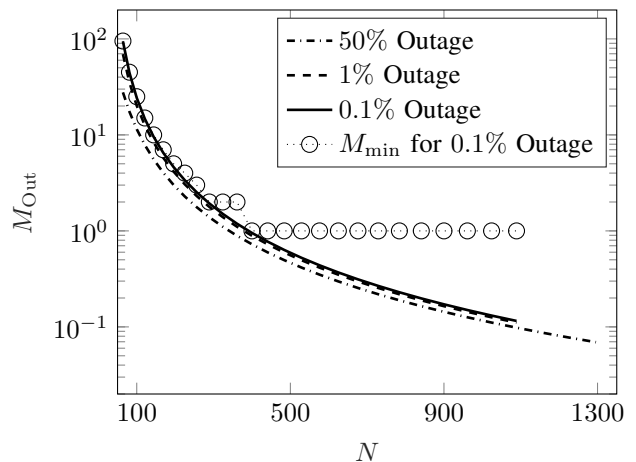
if $M \geq M_{\mathrm{Out}}$, where M_{Out} a positive real-valued solution to the following fixed-point equation in x:

$$\mathcal{R} + \frac{C}{1 + \rho \mu \mathrm{x}} \sqrt{B \mathrm{x}^2 + \omega \alpha_{\mathrm{d}} A_M \mathrm{x}} = \log_2 (1 + \rho \mu \mathrm{x}). \quad (58)$$

Here, μ and ω are defined for the given IRS as in Proposition 1. Moreover, C and B are given by

$$C = \rho Q^{-1} (p_{\mathrm{out}}) \log_2 \mathrm{e}, \quad (59a)$$
$$B = (\omega + 1) \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}} + \alpha_{\mathrm{d}} A_M. \quad (59b)$$

Fig. 11 demonstrates the IRS-BS dimensional trade-off with respect to the outage probability for various target outages. Here, the setting is set to be the same as the one considered in Fig. 10. The target rate is moreover set to $\mathcal{R} = 3$. For 50% outage, the trade-off curve recovers the one given in Fig. 10 for $\bar{\mathcal{C}} = 3$. This follows from the symmetry of the Gaussian density which leads to this property that the median coincides with the mean. We further observe that by reducing the outage probability, the trade-off figure shifts upward. This is intuitive, as for lower-outages, we require a better end-to-end link. These upward shifts are however not significant and rather negligible for large choices of N. This is a direct result of the channel hardening: With large N, the capacity term $\mathcal{C}$ is almost deterministic, and hence the outage probability tends to a step function in $\mathcal{R}$.

V. DERIVATION OF THE MAIN RESULT

This section proves the asymptotic result given in Proposition 1. The derivation follows three main steps:

- First, we derive the distribution of the end-to-end SNR gain Γ for an arbitrary M and N.
- We then bound the mean and variance of the SNR gain in terms of the spectrum of the covariance matrix $\mathbf{R}$.
- Finally, we send $N \to \infty$ and show that the maximum mutual information converges in distribution to a Gaussian random variable, if the maximum eigenvalue of $\mathbf{R}$ grows *sub-linearly* in N.

A. Distribution of End-to-End SNR

Theorem 1 determines the statistics of the end-to-end SNR gain Γ for an arbitrary choice of system dimensions.

Theorem 1. *Let the vector of phase-shifts at the IRS be set to $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_N]^\mathsf{T}$. Define $E(\cdot)$ as*

$$E(\boldsymbol{\beta}) = \bar{\alpha}_N \, \bar{\mathbf{g}}_\mathrm{r}^\mathsf{T} \boldsymbol{\Phi}(\boldsymbol{\beta}) \mathbf{R} \boldsymbol{\Phi}^\mathsf{H}(\boldsymbol{\beta}) \, \bar{\mathbf{g}}_\mathrm{r}^*, \qquad (60)$$

for matrix $\boldsymbol{\Phi}(\boldsymbol{\beta}) = \mathrm{diag}\left\{ \mathrm{e}^{-\mathrm{j}\beta_1}, \ldots, \mathrm{e}^{-\mathrm{j}\beta_N} \right\}$, $\bar{\alpha}_N$ given in (25) and $\bar{\mathbf{g}}_\mathrm{r} = \mathbf{a}_N(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1})$. Let $F_\zeta(\boldsymbol{\beta})$ and $\Lambda(\boldsymbol{\beta})$ be

$$F_\zeta(\boldsymbol{\beta}) = \bar{\alpha}_N \kappa_\mathrm{r} \left| \bar{\mathbf{g}}_\mathrm{r}^\mathsf{T} \boldsymbol{\Phi}(\boldsymbol{\beta}) \bar{\mathbf{h}}_\mathrm{r} \right|^2 + \zeta E(\boldsymbol{\beta}) \qquad (61\mathrm{a})$$

$$\Lambda(\boldsymbol{\beta}) = \frac{\alpha_\mathrm{d} A_M}{M} + E(\boldsymbol{\beta}). \qquad (61\mathrm{b})$$

Then, the end-to-end SNR gain Γ is distributed with a generalized chi-square distribution of order $2M$ whose density is given by

$$f_\Gamma(\gamma) = \frac{\tau_M}{\Lambda(\boldsymbol{\beta})} \int_{-\infty}^{+\infty} f_1\left(\frac{\gamma - v}{\Lambda(\boldsymbol{\beta})} \right) f_0(\tau_M v)\, \mathrm{d}v, \qquad (62)$$

for $\tau_M = M/\alpha_\mathrm{d} A_M$ and the density functions

$$f_0(v) = \frac{v^{M-2}}{(M-2)!} \mathrm{e}^{-v}, \qquad (63\mathrm{a})$$

$$f_1(v) = \mathrm{e}^{-v-\lambda(\boldsymbol{\beta})} I_0\left(2\sqrt{\lambda(\boldsymbol{\beta}) v} \right), \qquad (63\mathrm{b})$$

with $\lambda(\boldsymbol{\beta}) = F_0(\boldsymbol{\beta})/\Lambda(\boldsymbol{\beta})$, and denoting the modified Bessel function of the first kind and order zero. Moreover, the mean and variance of Γ are given by

$$\mu_\Gamma(\boldsymbol{\beta}) = \alpha_\mathrm{d} A_M + F_1(\boldsymbol{\beta}), \qquad (64)$$

and (65), given on the top of the next page, respectively.

Proof. Let $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_N]^\mathsf{T}$. For a given index $m \in [M]$, define $g_m(\boldsymbol{\beta})$ be defined as

$$g_m(\boldsymbol{\beta}) = \sum_{n=1}^{N} \mathrm{e}^{-\mathrm{j}\beta_n} h_{\mathrm{r},n} t_{nm}. \qquad (66)$$

Here, β_n and t_{nm} are deterministic scalars, and $h_{\mathrm{r},n}$ are jointly Gaussian for $n \in [N]$. We can hence conclude that $g_m(\boldsymbol{\beta})$ is distributed Gaussian with mean and variance

$$\mu_m(\boldsymbol{\beta}) = \frac{\mu_0(\boldsymbol{\beta})}{\sqrt{M}} \mathrm{e}^{-\mathrm{j}\Phi_m(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2})}, \qquad (67\mathrm{a})$$

$$\tilde{\sigma}^2(\boldsymbol{\beta}) = \bar{\alpha}_N \, \bar{\mathbf{g}}_\mathrm{r}^\mathsf{T} \boldsymbol{\Phi}(\boldsymbol{\beta}) \mathbf{R} \boldsymbol{\Phi}^\mathsf{H}(\boldsymbol{\beta}) \, \bar{\mathbf{g}}_\mathrm{r}^*, \qquad (67\mathrm{b})$$

respectively, where $\bar{\mathbf{g}}_\mathrm{r} = \mathbf{a}_N(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1})$, $\bar{\alpha}_N$ is defined in (25),

$$\mu_0(\boldsymbol{\beta}) = \sqrt{M \kappa_\mathrm{r} \bar{\alpha}_N} \, \bar{\mathbf{g}}_\mathrm{r}^\mathsf{T} \boldsymbol{\Phi}(\boldsymbol{\beta}) \bar{\mathbf{h}}_\mathrm{r}, \qquad (68)$$

and $\boldsymbol{\Phi}(\boldsymbol{\beta}) = \mathrm{diag}\left\{ \mathrm{e}^{-\mathrm{j}\beta_1}, \ldots, \mathrm{e}^{-\mathrm{j}\beta_N} \right\}$.

As h_m for $m \in [M]$ have common term $g_m(\boldsymbol{\beta})$, the entries of $\mathbf{h}$ are correlated. We hence determine the covariance of $\mathbf{h}$. To this end, we first note that

$$\boldsymbol{\mu}(\boldsymbol{\beta}) = [\mu_1(\boldsymbol{\beta}), \ldots, \mu_M(\boldsymbol{\beta})]^\mathsf{T} = \mu_0(\boldsymbol{\beta}) \mathbf{v}^* \qquad (69)$$

where $\mathbf{v} = \mathbf{a}_M(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2})/\sqrt{M}$. As the result, the covariance matrix is given by

$$\mathbf{C} = \mathcal{E}\left[(\mathbf{h} - \mu_0(\boldsymbol{\beta}) \mathbf{v}^*)(\mathbf{h}^\mathsf{H} - \mu_0^*(\boldsymbol{\beta}) \mathbf{v}^\mathsf{T}) \right], \qquad (70\mathrm{a})$$

$$= \alpha_\mathrm{d} A_M \mathbf{I}_M + M \tilde{\sigma}^2(\boldsymbol{\beta}) \mathbf{v}^* \mathbf{v}^\mathsf{T} = \mathbf{C}_0^2, \qquad (70\mathrm{b})$$

where $\mathbf{C}_0 = \sqrt{\alpha_\mathrm{d} A_M}\, \mathbf{I}_M + \varpi(\boldsymbol{\beta}) \mathbf{v}^* \mathbf{v}^\mathsf{T}$ with ϖ being

$$\varpi(\boldsymbol{\beta}) = \sqrt{\alpha_\mathrm{d} A_M + M \tilde{\sigma}^2(\boldsymbol{\beta})} - \sqrt{\alpha_\mathrm{d} A_M}. \qquad (71)$$

We now represent the end-to-end channel $\mathbf{h}$ as $\mathbf{h} = \mathbf{C}_0 \mathbf{h}_0$, where $\mathbf{h}_0 \sim \mathcal{CN}\left(\mathbf{C}_0^{-1} \boldsymbol{\mu}(\boldsymbol{\beta}), \mathbf{I}_M \right)$. Consequently, we have

$$\Gamma = \frac{1}{M} \mathbf{h}_0^\mathsf{H} \mathbf{C} \mathbf{h}_0. \qquad (72)$$

To determine this weighted norm, we use the eigenvalue decomposition of $\mathbf{C}$ to write $\mathbf{C} = \mathbf{V}^* \boldsymbol{\Sigma} \mathbf{V}^\mathsf{T}$, where $\mathbf{V} \in \mathbb{C}^{M \times M}$ is a unitary matrix whose *last* column is $\mathbf{v}$ and $\boldsymbol{\Sigma}$ is an $M \times M$ diagonal matrix defined as

$$\boldsymbol{\Sigma} = \mathrm{diag}\left\{ \alpha_\mathrm{d} A_M, \ldots, \alpha_\mathrm{d} A_M, \alpha_\mathrm{d} A_M + M \tilde{\sigma}^2(\boldsymbol{\beta}) \right\}. \quad (73)$$

By replacing into (72), we have

$$\Gamma = \frac{\alpha_\mathrm{d} A_M}{M} \sum_{m=1}^{M-1} |r_m|^2 + \frac{\alpha_\mathrm{d} A_M + M \tilde{\sigma}^2(\boldsymbol{\beta})}{M} |r_M|^2 \qquad (74)$$

where r_m is the m-th entry of $\mathbf{r} = \mathbf{V}^\mathsf{T} \mathbf{h}_0$. Noting that $\mathbf{V}^\mathsf{T}$ is unitary, we conclude that $\mathbf{r} \sim \mathcal{CN}\left(\boldsymbol{\mu}_\mathrm{r}(\boldsymbol{\beta}), \mathbf{I}_M \right)$, where

$$\boldsymbol{\mu}_\mathrm{r}(\boldsymbol{\beta}) = \mathbf{V}^\mathsf{T} \mathbf{C}_0^{-1} \boldsymbol{\mu}(\boldsymbol{\beta}). \qquad (75)$$

Since $\mathbf{C} = \mathbf{C}_0^2$, the matrix $\mathbf{C}_0^{-1}$ is decomposed as

$$\mathbf{C}_0^{-1} = \mathbf{V}^* \sqrt{\boldsymbol{\Sigma}}^{-1} \mathbf{V}^\mathsf{T}, \qquad (76)$$

where $\sqrt{\boldsymbol{\Sigma}}$ is a diagonal matrix whose first $M-1$ diagonal entries is $\sqrt{\alpha_\mathrm{d} A_M}$ and the last one is $\sqrt{\alpha_\mathrm{d} A_M} + \varpi(\boldsymbol{\beta})$. Hence, we can write

$$\boldsymbol{\mu}_\mathrm{r}(\boldsymbol{\beta}) = \mathbf{V}^\mathsf{T} \mathbf{V}^* \sqrt{\boldsymbol{\Sigma}}^{-1} \mathbf{V}^\mathsf{T} \boldsymbol{\mu}(\boldsymbol{\beta}) \qquad (77\mathrm{a})$$

$$= \mu_0(\boldsymbol{\beta}) \sqrt{\boldsymbol{\Sigma}}^{-1} \mathbf{V}^\mathsf{T} \mathbf{v}^*. \qquad (77\mathrm{b})$$

Noting that $\mathbf{v}$ is the last column of the *unitary* matrix $\mathbf{V}$, we conclude that $\mathbf{V}^\mathsf{T} \mathbf{v}^*$ is a standard base vector with the first $M-1$ being zero and the last entry being 1. This concludes that $\mu_{\mathrm{r},m}(\boldsymbol{\beta}) = 0$ for $m \in [M-1]$ and

$$\mu_{\mathrm{r},M}(\boldsymbol{\beta}) = \frac{\mu_0(\boldsymbol{\beta})}{\sqrt{\alpha_\mathrm{d} A_M} + \varpi(\boldsymbol{\beta})} \qquad (78\mathrm{a})$$

$$= \frac{\mu_0(\boldsymbol{\beta})}{\sqrt{\alpha_\mathrm{d} A_M + M \tilde{\sigma}^2(\boldsymbol{\beta})}} \qquad (78\mathrm{b})$$

We now define *independent* random variables

$$V_0 = \sum_{m=1}^{M-1} |r_m|^2, \qquad (79\mathrm{a})$$

$$V_1 = |r_M|^2. \qquad (79\mathrm{b})$$

By definition, V_0 is distributed chi-square with $2M-2$ degrees for freedom whose mean and variance are $M-1$ and whose density function is given by (63a). The random variable V_1 is further distributed non-central chi-square with two degrees of freedom and non-centrality parameter

$$\lambda(\boldsymbol{\beta}) = |\mu_{\mathrm{r},M}(\boldsymbol{\beta})|^2. \qquad (80)$$

Consequently, the mean and variance of V_1 are $1 + \lambda(\boldsymbol{\beta})$ and $1 + 2\lambda(\boldsymbol{\beta})$, respectively, and its probability density is given by (63b) in Theorem 1.

$$\sigma_\Gamma^2 \left(\boldsymbol{\beta} \right) = \frac{\alpha_\mathrm{d} A_M}{M} \left[2 F_1 \left(\boldsymbol{\beta} \right) + \alpha_\mathrm{d} A_M \right] + E \left(\boldsymbol{\beta} \right) \left[F_0 \left(\boldsymbol{\beta} \right) + F_1 \left(\boldsymbol{\beta} \right) \right]. \tag{65}$$

The SNR gain Γ is determined in terms of V_0 and V_1 as

$$\Gamma = \frac{\alpha_\mathrm{d} A_M}{M} V_0 + \frac{\alpha_\mathrm{d} A_M + M \tilde{\sigma}^2 \left(\boldsymbol{\beta} \right)}{M} V_1. \tag{81}$$

Noting that V_0 and V_1 are independent, the proof of Theorem 1 is concluded after defining the functions $E \left(\boldsymbol{\beta} \right) = \tilde{\sigma}^2 \left(\boldsymbol{\beta} \right)$ and $F_\zeta \left(\boldsymbol{\beta} \right) = \left| \mu_0 \left(\boldsymbol{\beta} \right) \right|^2 / M + \zeta E \left(\boldsymbol{\beta} \right)$. $\qquad \square$

B. Maximum Average End-to-End SNR

Using IRSs in the system, we are often interested in tuning the IRS phase-shifts, such that the average end-to-end SNR is maximized, i.e., $\mu_\Gamma \left(\boldsymbol{\beta} \right)$ is maximized over $\boldsymbol{\beta}$. Considering Theorem 1, the maximum mean SNR is not necessarily found in a closed-form. Nevertheless, closed-form lower and upper bounds can be derived in terms of the spectrum of $\mathbf{R}$.

Theorem 2. *Let the phase-shifts at the IRS be set to the vector $\boldsymbol{\beta}^\star = \left[\beta_1^\star, \ldots, \beta_N^\star \right]^\mathsf{T}$ with $\beta_n^\star$ being specified in (20). Then, the average end-to-end SNR gain is bounded as*

$$\mu_\Gamma \left(\boldsymbol{\beta}^\star \right) \leq \xi_\mathrm{max} \left(N \right) + \kappa_\mathrm{r} \bar{\alpha}_N N^2, \tag{82a}$$
$$\mu_\Gamma \left(\boldsymbol{\beta}^\star \right) \geq \xi_\mathrm{min} \left(N \right) + \kappa_\mathrm{r} \bar{\alpha}_N N^2 \tag{82b}$$

where $\xi_i \left(N \right)$ for $i \in \{ \mathrm{min}, \mathrm{max} \}$ is defined as

$$\xi_i \left(N \right) = \alpha_\mathrm{d} A_M + \bar{\alpha}_N \lambda_i N \tag{83}$$

with $\bar{\alpha}_N$ being given in (25), and λ_min and λ_max denoting the minimum and maximum eigenvalue of $\mathbf{R}$, respectively. The variance of Γ is further bounded as

$$P_\mathrm{min} \left(N \right) \leq \sigma_\Gamma^2 \left(\boldsymbol{\beta}^\star \right) \leq P_\mathrm{max} \left(N \right), \tag{84}$$

where $P_i \left(N \right)$ is given for $i \in \{ \mathrm{min}, \mathrm{max} \}$ in (85) at the top of the next page.

Proof. We start the proof by stating the following lemma:

Lemma 1. *For the function $F_0 \left(\boldsymbol{\beta} \right)$ defined in (61a), we have*

$$\max_{\beta_1, \ldots, \beta_N} F_0 \left(\boldsymbol{\beta} \right) = \kappa_\mathrm{r} \bar{\alpha}_N N^2, \tag{86}$$

being obtained by setting $\beta_n = \beta_n^\star$ given in (20) for $n \in [N]$.

Proof. The proof is given in the Appendix A. $\qquad \square$

By setting $\beta_n = \beta_n^\star$ in Theorem 1, the mean and variance of the SNR gain reduce to

$$\mu_\Gamma \left(\boldsymbol{\beta}^\star \right) = \alpha_\mathrm{d} A_M + E \left(\boldsymbol{\beta}^\star \right) + \kappa_\mathrm{r} \bar{\alpha}_N N^2, \tag{87}$$

and the expression given in (88) at the top of the next page, respectively. The mean can hence be bounded as

$$\mathcal{E}_\mathrm{min} + \kappa_\mathrm{r} \bar{\alpha}_N N^2 \leq \mu_\Gamma \left(\boldsymbol{\beta}^\star \right) \leq \mathcal{E}_\mathrm{max} + \kappa_\mathrm{r} \bar{\alpha}_N N^2, \tag{90}$$

and the variance reads

$$\sigma_{\Gamma, \mathrm{min}}^2 \leq \sigma_\Gamma^2 \left(\boldsymbol{\beta}^\star \right) \leq \sigma_{\Gamma, \mathrm{max}}^2, \tag{91}$$

with $\sigma_{\Gamma, i}^2$ being defined in (89) at the top of the next page for $i \in \{ \mathrm{min}, \mathrm{max} \}$. Here,

$$\mathcal{E}_\mathrm{max} = \alpha_\mathrm{d} A_M + \max_{\boldsymbol{\beta} \in \mathbb{R}^N} E \left(\boldsymbol{\beta} \right) \tag{92}$$
$$\mathcal{E}_\mathrm{min} = \alpha_\mathrm{d} A_M + \min_{\boldsymbol{\beta} \in \mathbb{R}^N} E \left(\boldsymbol{\beta} \right) \tag{93}$$

whose values are given by the following lemma:

Lemma 2. *For the function $E \left(\boldsymbol{\beta} \right)$ defined in (60), we have*

$$\xi_\mathrm{min} \left(N \right) \leq \alpha_\mathrm{d} A_M + E \left(\boldsymbol{\beta} \right) \leq \xi_\mathrm{max} \left(N \right) \tag{94}$$

for any $\boldsymbol{\beta} \in \mathbb{R}^N$, where the function $\xi_i \left(N \right)$ is defined in (83) for $i \in \{ \mathrm{min}, \mathrm{max} \}$.

Proof. The proof is given in the Appendix B. $\qquad \square$

By substituting the results of Lemma 2 into the bounds in (91) and (90), the proof is concluded. $\qquad \square$

Remark 2. *From the proof, it is straightforward to conclude that the upper-bounds given in Lemma 2 are in general valid for any choice of IRS phase-shifts. This means that, for any $\boldsymbol{\beta} \in \mathbb{R}^N$, we have*

$$\mu_\Gamma \left(\boldsymbol{\beta} \right) \leq \xi_\mathrm{max} \left(N \right) + \kappa_\mathrm{r} \bar{\alpha}_N N^2, \tag{95a}$$
$$\sigma_\Gamma^2 \left(\boldsymbol{\beta} \right) \leq P_\mathrm{max} \left(N \right). \tag{95b}$$

Remark 3. *Lemma 1 explains the logic behind the choice of IRS phase-shifts in Proposition 1. It is worth noting that the proposed phase-shifts in (20) do not necessarily maximize the average end-to-end SNR. Nevertheless, they are still good enough to guarantee the end-to-end channel hardening.*

C. Proof of Proposition 1

The proof of Proposition 1 follows readily from Theorems 1 and 2: Using Theorem 1, we represent the end-to-end SNR gain as $\Gamma = \mu_\Gamma \left(\boldsymbol{\beta}^\star \right) + \sigma_\Gamma \left(\boldsymbol{\beta}^\star \right) \tilde{\Gamma}$, where $\tilde{\Gamma}$ is a zero-mean and unit-variance generalized chi-square random variable denoting the centralized and normalized version of Γ. We now replace it in (18) and use the Taylor series of the logarithm to write

$$\mathcal{C} = \log_2 \left(1 + \rho M \mu_\Gamma \left(\boldsymbol{\beta}^\star \right) + \rho M \sigma_\Gamma \left(\boldsymbol{\beta}^\star \right) \tilde{\Gamma} \right), \tag{96a}$$
$$= \log_2 \left(1 + \rho M \mu_\Gamma \left(\boldsymbol{\beta}^\star \right) \right) + \log_2 \left(1 + \varsigma_N \tilde{\Gamma} \right) \tag{96b}$$
$$= \log_2 \left(1 + \rho M \mu_\Gamma \left(\boldsymbol{\beta}^\star \right) \right) + \varsigma_N \log_2 \mathrm{e} \, \tilde{\Gamma} + \epsilon_N, \tag{96c}$$

where ς_N is defined as[1]

$$\varsigma_N = \frac{\rho M \sigma_\Gamma \left(\boldsymbol{\beta}^\star \right)}{1 + \rho M \mu_\Gamma \left(\boldsymbol{\beta}^\star \right)}, \tag{97}$$

and ϵ_N is a polynomial in $\varsigma_N \tilde{\Gamma}$ satisfying

$$\epsilon_N = \mathcal{O} \left(\varsigma_N^2 \tilde{\Gamma}^2 \right). \tag{98}$$

---

[1]We use subscript N to indicate its dependency on dimension N.

$$P_i(N) = \frac{\alpha_{\mathrm{d}} A_M}{M} + \frac{2\alpha_{\mathrm{d}} A_M \lambda_i}{M} \bar{\alpha}_N N + \left(\frac{2\alpha_{\mathrm{d}} A_M \kappa_{\mathrm{r}}}{M} \bar{\alpha}_N + \bar{\alpha}_N^2 \lambda_i^2 \right) N^2 + 2\lambda_i \kappa_{\mathrm{r}} \bar{\alpha}_N^2 N^3 \tag{85}$$

$$\sigma_\Gamma^2(\boldsymbol{\beta}^\star) = \frac{\alpha_{\mathrm{d}} A_M}{M} \left[2\kappa_{\mathrm{r}} \bar{\alpha}_N N^2 + 2E(\boldsymbol{\beta}^\star) + \alpha_{\mathrm{d}} A_M \right] + E(\boldsymbol{\beta}^\star) \left[E(\boldsymbol{\beta}^\star) + 2\kappa_{\mathrm{r}} \bar{\alpha}_N N^2 \right], \tag{88}$$

$$\sigma_{\Gamma,i}^2 = \frac{\alpha_{\mathrm{d}} A_M}{M} \left(2\kappa_{\mathrm{r}} \bar{\alpha}_N N^2 + 2\mathcal{E}_i - \alpha_{\mathrm{d}} A_M \right) + \left(\mathcal{E}_i - \alpha_{\mathrm{d}} A_M \right) \left(\mathcal{E}_i + 2\kappa_{\mathrm{r}} \bar{\alpha}_N N^2 - \alpha_{\mathrm{d}} A_M \right), \tag{89}$$

Invoking Theorem 2, we can bound ς_N from above as

$$\varsigma_N \leq \frac{\rho M \sqrt{P_{\max}(N)}}{1 + \rho M \left(\xi_{\min}(N) + \kappa_{\mathrm{r}} \bar{\alpha}_N N^2 \right)} \tag{99a}$$

$$\overset{\dagger}{\leq} \frac{\rho M \sqrt{P_{\max}(N)}}{1 + \rho M \left(\alpha_{\mathrm{d}} A_M + \kappa_{\mathrm{r}} \bar{\alpha}_N N^2 \right)} \tag{99b}$$

where $\dagger$ follows from $\xi_{\min}(N) \geq \alpha_{\mathrm{d}} A_M$.

We now focus on scaling with N. Since $\lambda_{\max}$ grows *sublinearly* in N, we can conclude that there exist real scalars λ_0 and $0 \leq u < 1$, such that $\lambda_{\max}$ is uniformly bounded by

$$\lambda^{\mathrm{U}}(N) = \lambda_0 N^u, \tag{100}$$

From the constraint $A_{\mathrm{IRS}} = A_0 N^{1-q}$, we can further conclude that

$$A_N = \frac{A_0}{N^q}, \tag{101}$$

for some real A_0 and $0 \leq q < 1$, where the constraint in (22) guarantees that $u \geq q$. As a result, $P_{\max}(N)$ is uniformly bounded from above by

$$P^{\mathrm{U}}(N) = bN^{3+u-4q}. \tag{102}$$

These upper-bounds lead to this conclusion that ς_N is bounded as $\varsigma_N \leq \varsigma_N^{\mathrm{U}}$ for some ς_N^{U} satisfying

$$\varsigma_N^{\mathrm{U}} = \mathcal{O}\left(N^{\frac{u-1}{2}} \right). \tag{103}$$

We now note that $\tilde{\Gamma}$ is zero-mean and unit-variance, and $0 \leq u < 1$. Thus, $\epsilon_N \to 0$ as N grows large, and hence $\mathcal{C}$ is well approximated by the first two terms in (96c).

We now use the following lemma:

Lemma 3. *Let $F_0(\boldsymbol{\beta})/\Lambda(\boldsymbol{\beta})$ and $\sigma_\Gamma^2(\boldsymbol{\beta})$ grow large. Then, we have*

$$\frac{\Gamma - \mu_\Gamma(\boldsymbol{\beta})}{\sigma_\Gamma(\boldsymbol{\beta})} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1). \tag{104}$$

Proof. The proof is given in the Appendix C. □

We first note that $\sigma_\Gamma^2(\boldsymbol{\beta}^\star)$ grows large with N. Moreover, $F_0(\boldsymbol{\beta}^\star) = \bar{\alpha}_N \kappa_{\mathrm{r}} N^2$ and $\Lambda(\boldsymbol{\beta}^\star)$ is uniformly bounded from above with $\mathcal{O}\left(N^{1+u-2q} \right)$ for some $0 \leq u < 1$. This concludes that $F_0(\boldsymbol{\beta}^\star)/\Lambda(\boldsymbol{\beta}^\star)$ grows large as N increases. We hence use Lemma 3 and conclude that $\tilde{\Gamma}$ converges in distribution to a zero-mean and unit-variance Gaussian random variable, as N grows asymptotically large.

Finally by noting that $\bar{\mathbf{g}}_{\mathrm{r}}^{\mathsf{T}} \boldsymbol{\Phi}(\boldsymbol{\beta}^\star) = \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}}$, we can write

$$E(\boldsymbol{\beta}^\star) = \bar{\alpha}_N \bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}} \mathbf{R} \bar{\mathbf{h}}_{\mathrm{r}}. \tag{105}$$

Substituting into (96c), Proposition 1 is concluded after few lines of standard calculations.

D. Derivation of Pseudo-Proposition 1

From our numerical validations in Section III-D we heuristically conclude that $\mathcal{C}$ is still well-approximated by a Gaussian random variable in the linearly scaling settings. Nevertheless, the variance of $\mathcal{C}$ in this case is not tightly approximated by Proposition 1. From the proof of Proposition 1, the inaccuracy comes from the fact that the necessary condition for Lemma 3 is not satisfied and hence, $\tilde{\Gamma}$ does not converge in distribution to a zero-mean and unit-variance Gaussian random variable[1].

The exact distribution of $\tilde{\Gamma}$ in this case can be derived explicitly from Theorem 1. We however follow our heuristic and approximate it by a zero-mean Gaussian random variable of a smaller variance. Following the proof of Lemma 3, we show in Appendix D that $\tilde{\Gamma}$ can be approximated in this case as

$$\tilde{\Gamma} \approx \tilde{\Gamma}_\infty + \hat{\epsilon} \tag{106}$$

where $\tilde{\Gamma}_\infty \sim \mathcal{N}\left(0, \sigma_\infty^2 \right)$ with

$$\sigma_\infty^2 = \frac{\kappa_{\mathrm{r}}}{\kappa_{\mathrm{r}} + \vartheta} \tag{107}$$

and ϑ being given in (49). The random variable $\hat{\epsilon}$ is further a centralized[2] chi-square residual term. By replacing (106) in the Taylor expansion (96c) and ignoring the residual term as well as higher-order terms, the approximation is derived.

VI. CONCLUSIONS

Even with highly correlated reflecting elements, a large IRS hardens the end-to-end channel of a MIMO system with small transmit and receive array antennas. This conclusion follows from the main result of this study which shows that the mutual information between the input and output signals in an IRS-aided system is almost perfectly approximated by a Gaussian random variable, whose mean increases with the IRS size and whose variance vanishes to zero as the IRS size grows large.

The results of this study indicate that the channel hardening property is rather generic in IRS-aided systems and is easily achieved by small transmit and receive array antennas.

[1] This is why Proposition 1 excludes the strictly-linear case.
[2] That means $\hat{\epsilon}$ is zero-mean.

More interestingly, the phase-shifts of reflecting elements are not required to be updated frequently and should only be appropriately matched with the large-scale statistics of the wireless channels, i.e., AoAs andAoDs. These findings reveal this interesting fact: Even with IRSs of moderate size whose physical dimensions are rather significantly smaller than the wavelength, we can still harden the end-to-end MIMO channel between the transmitter and the receiver.

Earlier investigations in the literature have demonstrated the promising performance of massive MIMO systems in various respects. Nevertheless, the complexity of implementing such systems left this technology challenging for commercial uses. Along with earlier results in the literature, the result of this study indicate that using IRSs, the key features of the massive MIMO technology can be achieved with rather small end-to-end dimensions.

Although this work studies a basic setting, the analytical results can still be employed to investigate various fundamental properties of IRS-aided MIMO systems. An example is given in Section IV, where we demonstrate the dimensional trade-off between the IRS and the BS: The larger the IRS is, the smaller the BS needs to be in order to achieve a target performance. Similar investigations and further extensions of the results to more advanced settings, and under more realistic assumptions, are interesting study directions. They are however out of the scope of this particular paper and are left as potential directions for future work.

APPENDIX A
PROOF OF LEMMA 1

Considering the definition of $F_0(\boldsymbol{\beta})$, one can write

$$F_0(\boldsymbol{\beta}) = \bar{\alpha}_N \kappa_{\mathrm{r}} \left| \sum_{n=1}^{N} \mathrm{e}^{\mathrm{j}(\Pi_n(\varphi_{\mathrm{r}1},\theta_{\mathrm{r}1})+\Pi_n(\varphi_{\mathrm{t}1},\theta_{\mathrm{t}1})-\beta_n)} \right|^2 . \quad (108)$$

The expression on the right hand side includes the amplitude of a sum whose summands are complex numbers on the unit circle. As the result, the amplitude of the sum is maximized by setting all the summands in-phase. This is obtained by setting $\beta_n = \beta_n^\star$ for $\beta_n^\star$ given in (20). In this case, the sum adds to N and $F(\boldsymbol{\beta}^\star)$ is given by the expression given in Lemma 1.

APPENDIX B
PROOF OF LEMMA 2

Let $\boldsymbol{\psi}(\boldsymbol{\beta}) = \boldsymbol{\Phi}^{\mathsf{H}}(\boldsymbol{\beta})\mathbf{g}_{\mathrm{r}}^{*}$. Hence, $E(\boldsymbol{\beta})$ can be written as

$$E(\boldsymbol{\beta}) = \bar{\alpha}_N \, \boldsymbol{\psi}^{\mathsf{H}}(\boldsymbol{\beta})\mathbf{R}\boldsymbol{\psi}(\boldsymbol{\beta}) \quad (109)$$

Since every entry of $\boldsymbol{\psi}(\boldsymbol{\beta})$ lies on the unit circle, we can write $\|\boldsymbol{\psi}(\boldsymbol{\beta})\|^2 = N$ for any $\boldsymbol{\beta} \in \mathbb{R}^N$. Consequently, the function

$$Q(\boldsymbol{\beta}) = \frac{1}{N}\boldsymbol{\psi}^{\mathsf{H}}(\boldsymbol{\beta})\mathbf{R}\boldsymbol{\psi}(\boldsymbol{\beta}) \quad (110)$$

determines the *Rayleigh quotient* of $\mathbf{R}$ at $\boldsymbol{\psi}(\boldsymbol{\beta})$. We now use the fact that the Rayleigh quotient is bounded in terms of the eigenvalue of $\mathbf{R}$, as follows [43]

$$\lambda_{\min} \leq Q(\boldsymbol{\beta}) \leq \lambda_{\max}, \quad (111)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalue of $\mathbf{R}$, respectively. Substituting into (109), Lemma 2 is concluded.

APPENDIX C
PROOF OF LEMMA 3

We start the proof by considering (74). Using the definition of $\Lambda(\boldsymbol{\beta})$ in Theorem 1, we have

$$\Gamma = \frac{\alpha_{\mathrm{d}} A_M}{M} \sum_{m=1}^{M-1} |r_m|^2 + \Lambda(\boldsymbol{\beta})\,|r_M|^2 . \quad (112)$$

Here, r_m for $m \in [M-1]$ are independent complex Gaussian random variables with zero mean and unit variance, and r_M is a complex unit-variance Gaussian random variable independent of r_m for $m \in [M-1]$. The mean of r_M, i.e., $\mu_{\mathrm{r},M}(\boldsymbol{\beta})$, is given in (78b). Using the definitions given in Theorem 1, it can be shown that

$$|\mu_{\mathrm{r},M}(\boldsymbol{\beta})| = \sqrt{\frac{F_0(\boldsymbol{\beta})}{\Lambda(\boldsymbol{\beta})}}. \quad (113)$$

Let $r_M = \tilde{r}_M + \mu_{\mathrm{r},M}(\boldsymbol{\beta})$ where $\tilde{r}_M$ is the centralized form of r_M. We can hence write

$$\Gamma = T_0 + 2\Lambda(\boldsymbol{\beta})\,\Re\left\{\tilde{r}_M \mu_{\mathrm{r},M}^{*}(\boldsymbol{\beta})\right\} + |\mu_{\mathrm{r},M}(\boldsymbol{\beta})|^2 \quad (114)$$

where T_0 is defined as

$$T_0 = \frac{\alpha_{\mathrm{d}} A_M}{M} \sum_{m=1}^{M-1} |r_m|^2 + \Lambda(\boldsymbol{\beta})\,|\tilde{r}_M|^2 . \quad (115)$$

We now consider the normalized SNR gain, i.e.,

$$\hat{\Gamma} = \frac{\Gamma}{\sigma_{\Gamma}(\boldsymbol{\beta})} \quad (116)$$

By replacing Γ with the expression in (114), we have

$$\hat{\Gamma} = \Gamma_0 + \Gamma_1 + c \quad (117)$$

where Γ_0 and Γ_1 are random expressions defined as

$$\Gamma_0 = \frac{T_0}{\sigma_{\Gamma}(\boldsymbol{\beta})} \quad (118\mathrm{a})$$

$$\Gamma_1 = 2\frac{\Lambda(\boldsymbol{\beta})}{\sigma_{\Gamma}(\boldsymbol{\beta})}\Re\left\{\tilde{r}_M \mu_{\mathrm{r},M}^{*}(\boldsymbol{\beta})\right\} \quad (118\mathrm{b})$$

and c is a deterministic constant given by

$$c = \frac{|\mu_{\mathrm{r},M}(\boldsymbol{\beta})|^2}{\sigma_{\Gamma}(\boldsymbol{\beta})}. \quad (119)$$

Despite its complicated form of $\sigma_{\Gamma}^2(\boldsymbol{\beta})$, we can use the fact that $F_0(\boldsymbol{\beta}) \leq F_1(\boldsymbol{\beta})$ in Theorem 1, and bound $\sigma_{\Gamma}^2(\boldsymbol{\beta})$ as

$$2\Lambda(\boldsymbol{\beta})F_0(\boldsymbol{\beta}) \leq \sigma_{\Gamma}^2(\boldsymbol{\beta}) \leq 2\Lambda(\boldsymbol{\beta})F_1(\boldsymbol{\beta}) + \alpha_{\mathrm{d}}^2 A_M^2. \quad (120)$$

Noting that $\alpha_{\mathrm{d}}^2 A_M^2$ is fixed, we can conclude that growth of $\sigma_{\Gamma}^2(\boldsymbol{\beta})$ guarantees that $\Lambda(\boldsymbol{\beta})F_0(\boldsymbol{\beta})$ grows large.

We use the upper bound to write

$$\Gamma_0 \leq \frac{T_0}{2\sqrt{\Lambda(\boldsymbol{\beta})F_0(\boldsymbol{\beta})}} \quad (121\mathrm{a})$$

$$= \frac{\alpha_{\mathrm{d}} A_M \displaystyle\sum_{m=1}^{M-1} |r_m|^2}{2M\sqrt{\Lambda(\boldsymbol{\beta})F_0(\boldsymbol{\beta})}} + \sqrt{\frac{\Lambda(\boldsymbol{\beta})}{F_0(\boldsymbol{\beta})}}\,|\tilde{r}_M|^2 . \quad (121\mathrm{b})$$

Now, let $\sigma_\Gamma^2(\boldsymbol{\beta})$ and $F_0(\boldsymbol{\beta})/\Lambda(\boldsymbol{\beta})$ grow large. We can then conclude that the upper bound in this case converges to zero in the mean squared error (MSE). Since $\Gamma_0 \geq 0$, we can conclude that Γ_0 converges to zero.

The expression Γ_1 is further a real-valued zero-mean Gaussian random variable whose variance is given by

$$\sigma_1^2 = \frac{2\Lambda^2(\boldsymbol{\beta})}{\sigma_\Gamma^2(\boldsymbol{\beta})}|\mu_{\mathrm{r},M}(\boldsymbol{\beta})|^2. \qquad (122)$$

Using the lower and upper bounds on the variance of Γ, we can bound σ_1^2 as

$$\frac{F_0(\boldsymbol{\beta})}{F_1(\boldsymbol{\beta})} \leq \sigma_1^2 \leq 1. \qquad (123)$$

Noting that

$$F_1(\boldsymbol{\beta}) = F_0(\boldsymbol{\beta}) + \Lambda(\boldsymbol{\beta}) - \frac{\alpha_{\mathrm{d}}A_M}{M} \qquad (124)$$

we can further write

$$\lim_{\frac{F_0(\boldsymbol{\beta})}{\Lambda(\boldsymbol{\beta})} \to \infty} \frac{F_0(\boldsymbol{\beta})}{F_1(\boldsymbol{\beta})} = 1 \qquad (125)$$

which means that $\sigma_1^2 = 1$. As a result, by growth of $\sigma_\Gamma^2(\boldsymbol{\beta})$ and $F_0(\boldsymbol{\beta})/\Lambda(\boldsymbol{\beta})$, the normalized SNR gain $\hat{\Gamma}$ converges in distribution to a unit-variance real Gaussian random variable.

Finally, by noting that $(\Gamma - \mu_\Gamma)/\sigma_\Gamma$ is the centralized form of $\hat{\Gamma}$, we can conclude that

$$\frac{\Gamma - \mu_\Gamma(\boldsymbol{\beta})}{\sigma_\Gamma(\boldsymbol{\beta})} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1). \qquad (126)$$

This concludes the proof.

APPENDIX D
APPROXIMATED SNR FOR LINEAR SCALING

For a linearly scaling IRS area and $\lambda_{\max}$, the conditions in Lemma 3 are not satisfied, and hence the given limit is not valid. Considering the proof in Appendix C, this follows from the fact that due to linear scaling Γ_0 does not converge to zero, and accordingly the limit of σ_1^2 is not one. Considering (117) in Appendix C, we can write

$$\tilde{\Gamma} = \hat{\Gamma} - \frac{\mu_\Gamma(\boldsymbol{\beta})}{\sigma_\Gamma(\boldsymbol{\beta})} = \Gamma_1 + \hat{\epsilon}, \qquad (127)$$

where $\hat{\epsilon}$ is a centralized chi-square random variable[1]. The random variable Γ_1 is further zero-mean Gaussian whose variance is σ_1^2 given in (122).

Considering linear scaling, i.e., $q = u = 1$, $\mathbf{R}$ is a rank-one matrix and hence we can write $\mathbf{R} = \mathbf{r}\mathbf{r}^{\mathsf{H}}$ for some $\mathbf{r} \in \mathbb{C}^N$. As a result,

$$\bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}}\mathbf{R}\bar{\mathbf{h}}_{\mathrm{r}} = |\mathbf{h}^{\mathsf{H}}\mathbf{r}|^2. \qquad (128)$$

Noting that entries of $\mathbf{h}^{\mathsf{H}}$ lie equidistantly on the unit-circle, one can approximately write $\bar{\mathbf{h}}_{\mathrm{r}}^{\mathsf{H}}\mathbf{R}\bar{\mathbf{h}}_{\mathrm{r}} \approx 0$ for a typical[2] $\mathbf{r}$.

We further note that $\bar{\alpha}_N N^2$ converges to a constant with $q = 1$. Considering this fact and using the above approximation, the variance σ_1^2 is approximated with σ_∞^2 given in (107).

---

[1] $\hat{\epsilon}$ is zero-mean, since $\tilde{\Gamma}$ is zero-mean

[2] For instance, for the covariance matrix $\mathbf{R} = \mathbf{1}_N$, $\mathbf{r}$ is a vector of all-ones, and hence this approximation is almost exact.

REFERENCES

[1] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[2] D. Bai, P. Mitran, S. S. Ghassemzadeh, R. R. Miller, and V. Tarokh, "Rate of channel hardening of antenna selection diversity schemes and its implication on scheduling," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4353–4365, Sep. 2009.

[3] S. Asaad, A. M. Rabiei, and R. R. Müller, "Massive MIMO with antenna selection: Fundamental limits and applications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8502–8516, Nov. 2018.

[4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Nov. 2017.

[5] S. Asaad, A. Bereyhi, A. M. Rabiei, R. R. Müller, and R. F. Schaefer, "Optimal transmit antenna selection for massive MIMO wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 817–828, Apr. 2018.

[6] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 847–860, Apr. 2014.

[7] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[8] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[9] A. Bereyhi, S. Asaad, and R. R. Müller, "Stepwise transmit antenna selection in downlink massive multiuser MIMO," in *Proc. 22nd Inter. ITG Workshop on Smart Antennas*, Bochum, Germany, March, 2018, pp. 1–8.

[10] G. Caire, "On the ergodic rate lower bounds with applications to massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3258–3268, May 2018.

[11] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Stockholm, Sweden, Aug. 2015., pp. 201–205.

[12] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Net.*, vol. 2019, no. 1, pp. 1–20, May 2019.

[13] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2637–2652, Nov. 2020.

[14] A. Bereyhi, S. Asaad, R. R. Müller, R. F. Schaefer, and H. V. Poor, "Secure transmission in IRS-assisted MIMO systems with active eavesdroppers," in *Proc. IEEE 54th Asilomar Conf. Sig., Sys., Comp.*, Pacific Grove, CA, USA, Nov. 2020, pp. 718–725.

[15] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," in *Proc. IEEE 31st Int. Symp. Pers. Indoor Mob. Radio Commun. (PIMRC)*, London, UK, Sep., 2020., pp. 1–6.

[16] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[17] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation," *arXiv preprint arXiv:1906.02360*, 2019.

[18] H. Zhang, B. Di, Z. Han, H. V. Poor, and L. Song, "Reconfigurable intelligent surface assisted multi-user communications: How many reflective elements do we need?" *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1098–1102, May 2021.

[19] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2019.

[20] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surv*, vol. 22, no. 4, pp. 2283–2314, Jun. 2020.

[21] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, Jan. 2021.

[22] X. Li, J. Fang, F. Gao, and H. Li, "Joint active and passive beamforming for intelligent reflecting surface-assisted massive MIMO systems," *arXiv preprint arXiv:1912.00728*, 2019.

[23] Z. Wang, L. Liu, S. Zhang, and S. Cui, "Massive MIMO communication with intelligent reflecting surface," *arXiv preprint arXiv:2107.04255*, 2021.

[24] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for IRS-assisted full-duplex cognitive radio systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7376–7394, Sep. 2020.

[25] W. Wu, Z. Wang, L. Yuan, F. Zhou, F. Lang, B. Wang, and Q. Wu, "IRS-enhanced energy detection for spectrum sensing in cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2254–2258, Jul. 2021.

[26] Z. Ding and H. V. Poor, "A simple design of IRS-NOMA transmission," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1119–1123, Feb. 2020.

[27] Z. Ding, R. Schober, and H. V. Poor, "On the impact of phase shifting designs on IRS-NOMA," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1596–1600, Apr. 2020.

[28] Y. Tang, G. Ma, H. Xie, J. Xu, and X. Han, "Joint transmit and reflective beamforming design for IRS-assisted multiuser MISO SWIPT systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, July, 2020, pp. 1–6.

[29] S. Zargari, A. Khalili, and R. Zhang, "Energy efficiency maximization via joint active and passive beamforming design for multiuser MISO IRS-aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 557–561, Nov. 2020.

[30] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2052–2065, Jan. 2020.

[31] X. Zhang, X. Yu, and S. Song, "Outage and finite-SNR DMT analysis for IRS-aided MIMO systems: How large IRSs need to be?" *arXiv preprint arXiv:2111.15123*, 2021.

[32] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Dec. 2020.

[33] Z. Zhang, Y. Cui, F. Yang, and L. Ding, "Analysis and optimization of outage probability in multi-intelligent reflecting surface-assisted systems," *arXiv preprint arXiv:1909.02193*, 2019.

[34] H. Ibrahim, H. Tabassum, and U. T. Nguyen, "Exact coverage analysis of intelligent reflecting surfaces with Nakagami-m channels," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1072–1076, Dec. 2021.

[35] T. Wang, M.-A. Badiu, G. Chen, and J. P. Coon, "Performance analysis of IOS-Assisted NOMA system with channel correlation and phase errors," *arXiv preprint arXiv:2112.11512*, 2021.

[36] Z. Wang, L. Liu, and S. Cui, "Intelligent reflecting surface assisted massive MIMO communications," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Atlanta, GA, USA , May 2020, pp. 1–5.

[37] A. W. Van der Vaart, *Asymptotic statistics*. Cambridge university press, 2000, vol. 3.

[38] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Jun. 2019.

[39] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1663–1683, Jul. 2010.

[40] ——, "The multiport communication theory," *IEEE Circuits and Systems Magazine*, vol. 14, no. 3, pp. 27–44, Aug. 2014.

[41] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[42] S. Loyka, "The capacity of Gaussian MIMO channels under total and per-antenna power constraints," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1035–1043, Mar. 2017.

[43] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, New York, USA, 2013.

[44] I. The MathWorks, *Curve Fitting Toolbox*, Natick, Massachusetts, United State, 2020. [Online]. Available: https://www.mathworks.com/products/curvefitting.html

[45] H. Bölcskei, "MIMO-OFDM wireless systems: basics, perspectives, and challenges," *Proceedings of the IEEE*, vol. 13, no. 4, pp. 31–37, Feb. 2006.